DOCUMENT RESUME

ED 175 922 TM 009 564

AUTHOR Eddy, W. P.
TITLE Student Achievement in Alberta.
INSTITUTION Alberta Dept. of Education, Edmonton.
PUB DATE May 79
NOTE 97p.; Report prepared by the Minister's Advisory Committee on Student Achievement

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS *Academic Achievement; Achievement Tests; Attitude Tests; Educational Assessment; Educational Trends; Elementary Secondary Education; *Evaluation Needs; Foreign Countries; Performance Factors; *State Surveys; *Testing Programs; *Test Results
IDENTIFIERS *Alberta; Canada

ABSTRACT

The Minister's Advisory Committee on Student Achievement (MACOSA) of Alberta, Canada studied student achievement in grades 3, 6, 9, and 12. Students were tested in reading, writing, mathematics, science, and social studies (grade 12 only). Tests were based upon cognitive objectives considered to be important to the curriculum, and results were to be used for program improvement at the provincial level. Studies of factors affecting achievement and concerning the development of achievement and attitude tests were also conducted. Effects of the twelfth grade departmental examinations, which had been made non-compulsory, were investigated. General conclusions included: (1) third grade achievement levels were maintained over a 21-year period, from 1956 to 1977, and were comparable for urban and rural groups; (2) reading achievement was generally satisfactory at all four grade levels; (3) writing skills were weak on the short answer test, but higher on longer writing assignments; (4) mathematics skills were satisfactory in number facts and computation, but unsatisfactory in problem solving, geometry, measurement, and consumer mathematics; and (5) performance on the science tests was generally satisfactory. (Test data, a glossary and bibliography, and a copy of the order establishing MACOSA are included.) (GDC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED175922

TM009 564

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

W. P. Iddy

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Student Achievement in Alberta

May 1979

Alberta

Minister's Advisory Committee on Student Achievement

FOREWORD

The Minister's Advisory Committee on Student Achievement (MACOSA) was established in November 1976 to study the problems related to student achievement in Alberta and make recommendations for their solution. During its two-and-a-half-year term of office, MACOSA met for a total of 27 days and commissioned 18 studies. This report presents as briefly and clearly as possible (1) the major finds of the studies and (2) MACOSA's conclusions and recommendations, which are based to a large extent on these findings. Reports of each MACOSA study are also available for those who require more detailed information.

The members of MACOSA wish to acknowledge their debt to the many individuals and groups who made possible the successful completion of the project. Thanks are due to the organizations who sponsored the members of MACOSA; the researchers, steering committees and results interpretation panels; and the staff of Alberta Education who provided administrative and consultative services. Special thanks are extended to Beatrice Harke, Carol Warden, Sandy Forest and Cathy Stuart of Alberta Education for their efficiency in providing the clerical services essential to satisfactory completion of the project, frequently under stringent time constraints. Ruby Wilmer's special administrative assistance in the sampling procedures and on other projects is also recognized. Many others in school systems, universities and elsewhere made substantial contributions to MACOSA. Sincere thanks are extended to all those who helped in one way or another to carry out this demanding assignment.

Table of Contents

		Page
CHAPTER 1	**Background**	1
	Origins of MACOSA	1
	Membership	2
	Terms of Reference/Mandate	3
	List of Studies	6
	Procedures and Guidelines	7
CHAPTER 2	**Rationale For Evaluating Student Achievement**	10
	Definitions	10
	Purposes	11
	General Principles	11
CHAPTER 3	**Preliminary Studies**	13
	Communication Skills	13
	Social Studies	14
	Science	16
CHAPTER 4	**Studies on Factors Affecting Student Achievement**	18
CHAPTER 5	**Achievement Studies — Levels and Adequacy of Student Performance**	23
	Alberta Grade 3 Study	23
	Language Arts (Reading and Writing)	24
	Mathematics	33
	Science	39
	Social Studies	45
CHAPTER 6	**Test Development Studies**	52
	Listening and Speaking	52
	Attitudes Towards the World of Work	53
	Attitudes Towards School Subjects	55
	Assessment of Affective Objectives in the Social Studies	57
CHAPTER 7	**Grade 12 Departmental Examinations**	58
CHAPTER 8	**General Conclusions and Recommendations**	62
GLOSSARY		77

BIBLIOGRAPHY 79

APPENDIX A. **Ministerial Order Establishing MACOSA** 81

APPENDIX B. **The Goals of Basic Education for Alberta** 83

APPENDIX C. **Sources and Availability of MACOSA Studies** 85

APPENDIX D. **MACOSA Administrative and Consultative Staff** 85

APPENDIX E. **MACOSA Results Interpretation Panels** 86

APPENDIX F. **MACOSA Steering Committee Members** 89

List of Tables

		Page
Table 1	Reading Achievement Levels: Percentage of Correct Responses	26
Table 2	Writing Achievement Levels: Percentage of Correct Responses	29
Table 3	Longer Writing Assignment — Holistic Scoring	30
Table 4	Average Mathematics Achievement Levels for Content Strands: Percentage of Correct Responses	34
Table 5	Average Mathematics Achievement by Thought Level: Percentage of Correct Responses	35
Table 6	Science Achievement Levels: Percentage of Correct Responses	41
Table 7	Performance on Common Items on Science Tests: Percentage of Correct Responses	42
Table 8	Performance on Common Items on Science Tests by Content Area: Percentage of Correct Responses	42
Table 9	**STEP II Test Results: Average Raw Scores**	43
Table 10	TOUS, Form Jw, Test Results: Average Raw Scores	43
Table 11	Average Performance by Thought Level of Science Items: Percentage of Correct Responses	44
Table 12	Social Studies Achievement Levels: Average Raw Scores and Percentages	46
Table 13	Average Raw Scores for Grade 12 Students on Social Studies 10, 20 and 30 Items	47
Table 14	Specific Features of a Proposed Provincial Testing Program	67

List of Figures

		Page
Figure 1	Interaction Between Educational and Environmental Elements	19
Figure 2	Achievement Levels of Social Studies 10, 20 and 30 Students by Mental Process: Percentage of Correct Responses	49
Figure 3	Achievement Levels of Social Studies 10, 20 and 30 Students by Content Area: Percentage of Correct Responses	50
Figure 4	Achievement Levels of Social Studies 10, 20 and 30 Students on Canadian Content Items: Percentage of Correct Responses	51
Figure 5	Sample Item from Semantic Differential Test	56

CHAPTER 1

Origins

Background

In the early 1970's declining achievement scores among students in the United States, Great Britain and elsewhere received a great deal of publicity. For example, the American College Entrance Examination Board reported in 1976 that on the Scholastic Aptitude Test both verbal and mathematical scores had fallen steadily since 1963. A number of Albertans felt that the same situation existed in this province, and various interest groups discussed that possibility. Some staff members of Alberta universities and colleges claimed that high school graduates' academic skills were below acceptable levels, some parents expressed doubts about their children's mastery of basic skills, and various media reported an apparent decline in educational standards and a lack of a common approach in evaluation techniques.

Many Albertans became even more concerned after Minister of Education Hon. L.D. Hyndman's February 1973 announcement of the withdrawal of compulsory grade 12 departmental examinations. This decision, which shifted the major part of the responsibility for grade 12 student evaluation from the provincial to the local (school board) level, terminated an evaluation practice used for over 60 years. Shortly after the Minister's announcement, Alberta Education declared that all school jurisdictions in the province could become accredited[1] for the purpose of assigning grade 12 marks by meeting certain standards such as adherence to prescribed curricula and the employment of properly certificated teachers. The departmental examinations remained available for students in non-accredited school systems and for special purposes, including appeals and scholarship applications, but by 1974 all schools in Alberta had begun to assign their own grade 12 marks.

Universities, colleges, and members of the general public continued to ask questions about the acceptability of "accredited" final marks for grade 12 academic subjects and the adequacy of levels of achievement. Alberta Education formed several committees and commissioned reports to review these concerns. Two reports were published in 1975: *High School Diploma Requirements* (Warren) and *Evaluating Education in Alberta Schools* (Rusnack). Also in 1975, Dr. Bryant Stringham of the Field Services Branch of Alberta Education prepared an unofficial report entitled *Next Steps in Accreditation*. In 1976 the Alberta Teachers' Association issued a policy statement on accreditation and testing; the Alberta School Trustees' Association proposed an Educational Review Agency; and the Committee on the Articulation of High Schools and Post-Secondary Educational Institutions recommended an "assessment of the quality of education" in the form of a review of accreditation of Alberta high schools and admission standards for post-

[1] A glossary of educational terms appears at the end of this report.

secondary educational institutions. More specifically, the Articulation Committee proposed that such a review would determine whether or not comparable marks were being assigned to students and whether there was a need to establish provincial standards of achievement by means of Alberta-normed tests of basic skills.

Public debates on student achievement and evaluation procedures, as they applied to all levels, occurred more and more frequently but produced more heat than light because of the absence of reliable information upon which to base any practical solutions. In the midst of all this controversy Hon. Julian Koziak, Minister of Education, established the Ministers' Advisory Committee on Student Achievement (MACOSA) in October 1976 and commissioned it with the task of studying the problems surrounding this issue and making recommendations for their solution. At that time Mr. Koziak said:

> The quality of education in Alberta, and indeed in all of Canada, is one of the most controversial issues in education today.... The Committee findings should help us to make sound educational decisions based on facts and to eliminate myths and conjectures.

Membership

During its term of office, from November 1, 1976 to March 31, 1979[2], MACOSA held 20 meetings which involved a total meeting time of 27 days. The following people served as committee members:

Dr. James Hrabi (Chairman), Alberta Education

Mr. Dennis Cambly, Alberta Chamber of Commerce
- Mr. John Moxon from July 1977

Dr. N. P. Hrynyk, Alberta Teachers' Association
- Mr. Keith Harrison from August 1977 to September 1978
- Mr. W. M. Brooks from September 1978

Dr. Robert Lawson, Committee on the Articulation of High Schools and Post-Secondary Educational Institutions

Mr. Eugene Mitchell, Alberta Federation of Labour

Mr. Andrew Stojak, Alberta Advanced Education and Manpower

Mr. Michael Strembitsky, Conference of Alberta School Superintendents
- Dr. Harry Mosychuk represented Mr. Strembitsky.

Mr. Peter Swann, Alberta Federation of Home and School Associations
- Mrs. Joyce Westerlund from November 1978

Mr. Lawrence Tymko, Alberta School Trustees' Association
- Mr. Stan Maertz from September 1978

[2] Originally MACOSA's term of office was to end on December 31, 1978, but the Minister later extended the term to March 31, 1979.

MACOSA's task would have been impossible without the support and interest of the nominating groups and institutions. Their contribution is gratefully acknowledged.

Dr. H. G. Sherk, Associate Director of the Planning and Research Branch of Alberta Education, served as Executive Secretary and Project Director of MACOSA. Several other Alberta Education staff members, whose names appear in Appendix D, served in various administrative, consultative and technical capacities.

Terms of Reference/Mandate

MACOSA was charged with investigating levels of student achievement in Alberta schools and considering evaluation methods by performing certain tasks as set out in the Ministerial Order.[3] These tasks included:

1. Studying the elements of student achievement (definitions, purposes, principles, and underlying assumptions).
2. Considering ways and means for assessing student achievement.
3. Considering ways and means for maintaining and/or improving current levels of achievement.
4. Studying how the withdrawal of compulsory grade 12 departmental examinations has affected the quality of education.
5. Recommending appropriate structures and procedures for monitoring student achievement in the future.

MACOSA developed seven key questions related to its task. These questions served as an outline for this report:

1. How does MACOSA define "student achievement" and "evaluation of student achievement"? (Chapter 2)
2. What are the reasons for evaluating student achievement? (Chapter 2)
3. What factors affect student achievement? (Chapter 4)
4. What are the levels of student achievement in Alberta in selected subjects and grades? (Chapter 5)
5. How adequate are these levels? (Chapter 5)
6. What effect has the non-compulsory nature of grade 12 departmental examinations had on the quality of education? (Chapter 7)
7. What ways and means are available or needed for assessing and monitoring in order to improve student achievement? (Chapters 3 and 6 describe various studies that MACOSA conducted on ways and means of assessment, and Chapter 8 summarizes MACOSA's conclusions and recommendations.)

[3] The Ministerial Order appears in full in Appendix A.

MACOSA's review of instruction and achievement in Alberta schools excluded duplication of the efforts of other boards and committees. MACOSA did not consider problems related to finance, which were examined recently by the Minister's Advisory Committee on School Finance. Similarly, the work of the Curriculum Policies Board on curricular objectives, resource materials and teaching methods was excluded. To facilitate test development, MACOSA's researchers and steering committees found it necessary to establish priorities among the stated curricular objectives and in some cases to make the objectives more specific. On the whole, though, the prescribed Alberta curriculum, including the *Goals of Basic Education*[4], was accepted as given. After lengthy deliberation MACOSA decided as well that although information about relationships between high school and post-secondary achievement would be relevant to its mandate, the gathering and analyzing of data on this subject would fall more appropriately in the domain of such groups as the Committee on the Articulation of High Schools and Post-Secondary Educational Institutions. Consequently, MACOSA did not initiate any studies in this regard but did review the one study[5] on the subject which became available during its term of office.

After the mandate and terms of reference had been clarified, MACOSA planned a number of projects and studies. In late 1976 MACOSA established as its first priority a thorough study of achievement in several knowledge and skills areas. The achievement areas to be reviewed were drawn from the following goals in the 1975 interim edition of the *Goals of Basic Education:*

To develop skills in communication (listening, speaking, reading, writing, viewing).

a. Develop skill in understanding the communication of others.
b. Develop ability in communicating ideas and feelings effectively.
c. Develop skill in oral and written languages.

To develop basic and special knowledge competencies.

a. Develop understanding and skills in the use of numbers, natural sciences, mathematics and social sciences.
b. Develop a fund of information and concepts.
c. Develop special interests and abilities.

Specifically, MACOSA commissioned the following achievement studies: reading, language and arithmetic achievement of grade 3 students in Alberta, comparing mid-1950's levels to 1977 levels; and reading, writing, mathematics, and

[4] See Appendix B for details of the 1978 Alberta *Goals of Basic Education*. Before May of 1978 MACOSA used the 1975 interim edition of this goals statement, which was not substantially different.

[5] This study, entitled *High School Grades and University Performance* (Mehra, 1977), is described briefly in Chapter 7.

science achievement at the grades 3, 6, 9, and 12 levels in 1978. MACOSA also reported and interpreted the achievement levels of Social Studies 30 students who wrote Alberta Education's High School Achievement Test in June 1977.

A second major task resulted from a motion of the Legislative Assembly on October 19, 1976 requesting that a study be conducted to determine the effect of the non-compulsory nature of the grade 12 departmental examinations on the quality of education. MACOSA responded to this request by commissioning the *Alberta Grade 12 Examination Study*.

MACOSA's third major activity was a thorough review of the attitudinal objectives of the Alberta curriculum and of possible means for assessing achievement of these objectives. In 1978, noting the importance that the Alberta Education goals statement had placed on students' attitudes, MACOSA commissioned two studies to develop and begin validating tests which would measure attitudes towards school subjects and the world of work. These studies were based on the following goals of schooling, which appear in the Alberta *Goals of Basic Education*, 1978:

> To acquire basic knowledge and develop skills and attitudes in mathematics, the practical and fine arts, the sciences, and the social studies (including history and geography), with appropriate local, national, and international emphasis in each.
>
> To acquire knowledge and develop skills, attitudes, and habits required to respond to the opportunities and expectations of the world of work.

In the summer of 1978 MACOSA commissioned a third study related to values and attitudes. Because the Alberta social studies curriculum places special emphasis on the study of values, MACOSA asked that a researcher develop test models for assessing achievement of the affective goals of the social studies program.

The MACOSA language arts achievement study dealt only with reading and writing skills whereas the Alberta language arts curriculum also includes listening, speaking and viewing. To avoid undue emphasis on reading and writing skills, MACOSA commissioned a study to develop and validate tests of listening and speaking for grades 3, 6, 9 and 12.

The above-mentioned test development projects, both in the attitudinal areas and in listening and speaking, have broken new ground. No tests existed previously which were valid for the Alberta curriculum. Further refinement is required because the researchers had only enough time to develop the tests and begin validation procedures, but both the test developers and MACOSA are optimistic that these tests will be found valid and reliable and that they can serve a useful function in Alberta schools.

MACOSA also considered the numerous educational and environmental factors which might influence student achievement. Two reviews of the literature were

commissioned — one focusing on school-based factors such as curriculum and teaching methods and a second on non-school factors such as home and community background. MACOSA recognized the importance of such factors, particularly in evaluating individual students. However, because of time constraints, the complexity of the subject and the scope of its mandate, MACOSA undertook no further work in this area.

Commissioned Studies

To provide the necessary information for drawing conclusions and making recommendations about student achievement, MACOSA commissioned a number of studies.[6] Smaller studies were commissioned directly, while larger ones required ministerial approval and supervision by steering committees. The steering committee members, whose names appear in Appendix F, were responsible for developing a study proposal, ensuring that project deadlines were met and progress reports were filed, providing technical assistance and advice on policy matters to researchers, supervising publication of the final report, and making recommendations in relation to the completed study.

All MACOSA's major studies[7] are listed below under three headings: (1) preliminary studies, (2) achievement studies, and (3) other studies.

Preliminary Studies
(Completed March - July, 1977)

1. *Assessing Communication Skills: A Review of the Literature* by N. Marchak and others (Edmonton Public School District).
2. *Ways and Means: Evaluation of Communication Skills* by G. Bevan, J. Gajadharsingh, and B. Gommeringer.
3. *Ways and Means: Evaluation of Listening and Speaking Skills* by B. Gommeringer, P. McFetridge, and E. Plattor.
4. *Feasibility Study on Assessing Student Achievement in Social Studies* by B. Chandler and others.
5. *Feasibility Study on Assessing Student Achievement in Science* by M. R. Treasure (Alberta Education).

Achievement Studies
(Completed May - October, 1978)

1. *Alberta Grade 3 Achievement Study* by S. C. T. Clarke, V. Nyberg, and W. Worth (Alberta Advisory Committee for Educational Studies).
2. *Alberta Language Arts (Reading and Writing) Achievement Study* by the Canadian Institute for Research in the Behavioral and Social Sciences.

[6] Besides these studies which were officially commissioned, many working papers, reports, steering committee minutes and other background materials (prepared by MACOSA's administrative and consultative staff members and others) were submitted to MACOSA as information.

[7] Appendix C provides information about distribution and sources of these studies.

3. *Alberta Assessment of School Mathematics* by A. Olson, D. Sawada, and S. Sigurdson (University of Alberta).
4. *Alberta Science Achievement Study* by M. R. Treasure (Alberta Education).
5. *Report on the Development and Administration of the High School Social Studies Achievement Tests* by R. Lopatka (Student Evaluation and Data Processing Services Branch, Alberta Education).

Other Studies

1. *Alberta Grade 12 Examination Study* by F. Dumont (Inter-Disc Consultants). Completed October 1977.
2. *School-Based Elements Related to Achievement: A Review of the Literature* by L. Fedigan. Completed September 1978.
3. *Elements Related to Student Success in Schooling and Education: A Review of the Literature* by G. Gay. Completed September 1978.
4. *Listening and Speaking Objectives* by Edmonton Public School District. Completed January 1978.
5. *Test Development for Assessing Achievement in Listening and Speaking* by E. Plattor, W. Unruh and L. Muir (University of Calgary). Completed October 1978.
6. *Development of Scales on Attitudes Towards the World of Work* by T. Maguire and E. Romaniuk (University of Alberta). Completed November 1978.
7. *Development of School Subjects Attitude Scales* by S. C. T. Clarke and V. Nyberg (University of Alberta). Completed December 1978.
8. *Development of Models for Assessing Affective Objectives in the Social Studies* by R. Lopatka. Completed December 1978.

Procedures and Guidelines

Within the above-mentioned terms of reference MACOSA carefully addressed the question of student achievement and drew together data that is as accurate and complete as possible. The researchers made every effort to relate their tests to the Alberta curriculum and to concentrate on testing knowledge, skill, and attitude objectives which are significant within the total curriculum. (The grade 12 tests in reading, writing, mathematics and science, however, were also designed to assess achievement of certain general objectives which, although they may not be specifically stated in the official grade 12 curriculum, are judged to reflect important school leaving competencies.) The researchers, with the assistance of the Curriculum Branch of Alberta Education and the steering committees, identified specific, measurable objectives and established priorities among the objectives before developing the test items.

The researchers followed standard test development procedures in each study. They piloted and field tested to check the reliability and validity of test items and deleted or modified inappropriate items. The sampling techniques were approximately the

same for all studies and were suitable for generalizing about the performance levels of the total school population in Alberta.

To encourage a high level of cooperation from the teachers and students involved, MACOSA explained the purposes of the testing project in letters sent to superintendents and school principals, who were requested to pass on the necessary information to schools. For the same reason MACOSA avoided overtesting any one group of students. For example, the mathematics achievement study used a sampling technique in which no student was required to answer all 120 items in the total test.

MACOSA decided that available information should be utilized as much as possible. As a preliminary project, MACOSA examined the work other provinces had done in the area of evaluating student achievement. British Columbia's Learning Assessment Project proved to be particularly useful in the development of suitable approaches for Alberta. MACOSA also benefitted from work done by the various branches of Alberta Education. The *Goals of Basic Education* statement, which was prepared by the Curriculum Policies Board, served as a guideline for selecting subjects to be researched. Information about sampling and test development techniques used by the Student Evaluation and Data Processing Services Branch was of great assistance. MACOSA encouraged the test developers to use existing tests or test items where feasible, but in most cases too few of the items were judged to be consistent with Alberta's curricular objectives to make existing tests acceptable.

Guidelines were also developed regarding communication, security of test items, and the release of test results for individuals or schools. An informal communication system included occasional progress reports, a news release prepared when the high-interest grade 12 departmental examination study was completed, and addresses by the Minister, members of MACOSA, or Alberta Education staff given to various groups interested in the Advisory Committee's projects. When public concerns became evident or questions arose, MACOSA responded by attempting to clear up misunderstandings and provide the necessary information.

MACOSA adopted a policy that security be maintained on complete test sets but that up to 40 percent of the individual items could be revealed in the reports as samples for clarification purposes. This decision resolved the problem of reporting test results without identifying so many of the actual questions used that the tests could not be administered in future.

MACOSA decided to publish scores and response patterns for test items only at the provincial level and not to report individual student and school scores either in MACOSA publications or to particular students and schools. A letter sent by the Executive Secretary of MACOSA to school principals outlined the reasons for this decision:

MACOSA is employing a stratified random sampling technique that economically produces valid information for the province as a whole. However, the schools selected within a specific school jurisdiction may not be typical for the jurisdiction. Consequently, the school results are useful only for provincial, not local, comparisons. Second, some MACOSA studies involve the technique of "matrix sampling", which enables use of a relatively small number of students for testing of a large number of curricular topics. Each subject test has several forms, and individual students in a school may be answering different questions which prevent local comparison of results because there is no common test. However, on the provincial level the various test forms can be analyzed relative to specific curricular topics which together provide a considerable amount of information about provincial achievement levels.

Another procedural decision arose from the fact that the data from the achievement studies provided only some of the information that MACOSA required, particularly with regard to the question, "Are the levels of student achievement at, below or above acceptable levels?" Because the tests had been administered only once, little information was available for judging the adequacy of performance or making long-term comparisons of student achievement. MACOSA decided that informed judgements should be sought from knowledgeable educators and members of the public by means of "results interpretation panels". These panels consisted of representatives of basic and post-secondary education, professional groups, labor, business, parents and school boards. Panel members, whose names appear in Appendix E, were invited to comment on how well they thought the students had performed. The researchers, the steering committees for each achievement study, and members of MACOSA also made such judgements. The statements in this report regarding adequacy of achievement (Chapter 5) represent the views of the Advisory Committee after considering the reports of researchers, results interpretation panels, and steering committees along with all other available data.

CHAPTER 2

Rationale for Evaluating Student Achievement

This chapter, which is intended to provide a rationale for MACOSA's evaluation activities, includes definitions of "student achievement" and "evaluation of student achievement", a set of purposes for evaluation, and a description of the general principles which served as guidelines.

Definitions

> Student achievement is defined as the level of performance of groups of Alberta students on tests based on selected objectives of schooling in certain subject areas.

For MACOSA's purposes student achievement is defined in terms of the results of tests administered at the grades 3, 6, 9, and 12 levels in reading, writing, mathematics, and science, and at the grade 12 level in social studies.

MACOSA's definition of student achievement is further qualified by the following characteristics of its achievement tests:

1. The achievement tests were developed to address a relatively few high priority objectives.
2. Cognitive objectives predominated. (Only objectives measurable by means of pencil and paper tests could be included.)
3. The objectives selected for the grades 3, 6, and 9 tests in all subjects and for the grade 12 social studies test were judged to be relevant to the Alberta curriculum. The more general objectives selected for the grade 12 tests in reading, writing, mathematics, and science, while judged to reflect important school leaving competencies, were not necessarily keyed to the curriculum.
4. Since the test exercises were based on selected objectives which represented high priority aspects of the curriculum, the results in most of the assessment studies were processed, interpreted and reported by objectives or clusters of objectives and not by total test scores. For example, the MACOSA reading achievement study shows how well students performed in obtaining the meanings of words from the context or in comprehending the main ideas of a passage, but it does not give a total test score for a particular grade level. The tests were keyed to objectives or clusters of objectives and they should be interpreted in those terms.

In summary, MACOSA defined student achievement as the level of performance of groups of Alberta students in achieving selected objectives of certain school programs. The types of achievement measured depended on what objectives were selected for inclusion in the tests; achievement levels in other parts of the school program did not receive attention.

MACOSA defined "evaluation of student achievement" in terms of the procedures used and the purposes served.

Evaluation of student achievement involves the collecting, analyzing, interpreting and reporting of appropriate information in summary form for use in making decisions on the conduct of public education. The information is related to goals and interpreted in terms of standards.

The information that MACOSA collected was delimited in the following ways:

1. The information gathered was intended primarily for use by those who have responsibilities and interests at the provincial level. Information directly related to achievement in individual schools and school systems was not collected.
2. Students and their parents did not receive feedback about individual student performance levels. (See letter to school principals, page 9).
3. MACOSA set standards after receiving the judgements of results interpretation panels and steering committees as to adequacy of achievement. In future the 1978 results, which provide benchmarks, could assist in setting standards.

Purposes

MACOSA determined that the following purposes were appropriate in terms of the Advisory Committee's mandate and resources.

MACOSA's evaluation activities were conducted for the purpose of monitoring and reporting information to the various publics as a basis for decision making about:

1) the continuation, discontinuation or modification of a program, and

2) the improvement of instruction for groups (program evaluation).

The MACOSA studies will provide decision makers at the provincial level with preliminary information regarding possible program modifications in areas where student performance is judged to be weak. The information will also be useful as a basis for making decisions about evaluation policies.

The purpose, "decision making about the improvement of instruction for groups" has been partially achieved, but only in a preliminary way and only for provincial programs. The information gathered was restricted to indicators of student performance and did not include processes or methods used to achieve the results. Therefore, the MACOSA reports were unable to provide definitive information about reasons for high or low performances. Many of the reports do include some speculation, though, which could provide guidance for appropriate action and further study.

General Principles

In planning and conducting its achievement studies, MACOSA considered definitions of and purposes for evaluation as described above, along with accepted evaluation principles. On that basis, the following guidelines were developed:

1. The information in the various MACOSA reports would be directed primarily to those who have responsibilities and interests at the provincial level.

2. MACOSA would use established procedures in conducting the achievement studies. In other studies, such as the development of listening and speaking tests, some new approaches would be used as required but only on a tentative basis. Provision would be made for further study before these new approaches were widely implemented.
3. All MACOSA achievement studies would use sampling by schools, and matrix sampling procedures would be employed in some studies.
4. MACOSA studies would consider a variety of reporting categories and relate each category to an information need. For example, some MACOSA studies might compare the achievement of boys to that of girls, and other studies might report achievement according to size of school. Information about socio-economic status, academic aptitude or motivation is not required for reporting on standards of achievement for groups, and therefore would not be gathered.
5. MACOSA's judgements about the adequacy of current performance would of necessity be based on viewpoints of various panels and committees. (The 1978 performance levels on MACOSA tests would provide benchmarks against which future achievement levels could be compared.)
6. MACOSA would attempt to ensure adequate communication between the evaluators and relevant audiences and publics. As an outgrowth of this decision MACOSA asked for public input in some of its preliminary studies, in the *Alberta Grade 12 Examination Study*, and through the use of the results interpretation panels which helped to establish standards of performance.
7. It was expected that the public would receive information about achievement levels through release of the reports by the Minister.

CHAPTER 3

Preliminary Studies

MACOSA's first step was to commission several preliminary studies to identify and evaluate ways and means for conducting province-wide assessments of communication skills, social studies and science.

Communication Skills

MACOSA conducted three preliminary studies related to communication skills: a review of the literature, ways and means for evaluating communication skills in general, and ways and means for evaluating listening and speaking.

1. Assessing Communication Skills: A Review of the Literature

Purpose The purpose of the study was to provide background information for developing appropriate tests of communication skills and generating reliable and valid data on achievement levels.

Design The researchers collected and summarized pertinent literature, noting in particular promising new assessment techniques.

Findings The study divided communication skills into five components: reading, writing, functional literacy, listening and speaking. Researchers have done the most work on the assessment of reading because traditional standardized reading tests have major limitations. Little work has been done on the assessment of writing, but researchers believe that tests should include actual writing activities. The study of functional literacy is relatively recent, and adult literacy testing is still a developing field marked by tests of varying quality. Studies of oral communication testing have been hampered by the complexity of procedures required for analysis, the degree of subjectivity involved, the availability of time, and the lack of qualified classroom personnel.

Conclusions and Observations New reading assessment techniques which would allow the measurement of more complex reading skills are becoming available. Increased computer assistance could greatly facilitate advances in this area. Valid assessments of writing in the future should be more comprehensive than they have been in the past and should include actual writing activities as well as other types of questions. Assessing functional literacy along with the more academic reading and writing skills would increase the comprehensiveness and validity of any large-scale literacy assessment project.

Oral communication tests developed in future will likely concentrate on measuring originality in form and content. The most successful testing of oral communication skills to date has taken place on an individual or small group basis and in a controlled but natural testing situation. Listening skills are more and more often being considered an important part of the ability to communicate, but further test development is needed in both listening and speaking.

Action by MACOSA On the basis of information provided by this study, MACOSA commissioned three subsequent studies — the language arts (reading and

writing) achievement study and two studies which identified objectives and developed tests for listening and speaking.[8] The reading and writing study is described in Chapter 5 and the listening and speaking studies are described in Chapter 6.

2. *Ways and Means: Evaluation of Communication Skills*

Purpose The purpose of the study was to develop a proposal for ways and means of assessing communication skills and knowledge among Alberta students.

Procedures The researchers gathered information from groups of Alberta educators, outlined the scope and sequence of the Alberta language arts program, and identified components of the program which are suitable for testing. They also reviewed available tests, suggested appropriate sampling techniques and prepared preliminary cost estimates.

Findings The study showed that almost no suitable tests are available, except for a few in the area of reading. The researchers concluded that an essay test should be developed in order to assess students' ability to use language rather than their ability to recognize errors. They noted that such a test could be developed in a relatively short time, but that costs for marking would be high. They also suggested further study of the desirability and feasibility of testing listening and speaking skills.

Action by MACOSA In response to the results of this study, MACOSA commissioned the reading and writing achievement study (Chapter 5) and the ways and means study of listening and speaking.

3. *Ways and Means: Evaluation of Listening and Speaking Skills*

Purpose The purpose of the study was to develop a proposal for ways and means of assessing listening and speaking skills among Alberta students.

Procedures The researchers developed a rationale for assessing listening and speaking, outlined the required tasks for carrying out an assessment program, suggested sampling procedures and prepared cost estimates.

Findings and Conclusions The researchers reported that an assessment of listening and speaking skills would be both desirable and feasible. They also made the following suggestions: professional educators should be involved in the planning stages; testable objectives should be identified before selecting and/or developing suitable tests; and students' speeches should be recorded on videotapes and used to train scorers.

Action by MACOSA In response to the results of the study MACOSA conducted two subsequent studies on listening and speaking, which are described in Chapter 6.

Social Studies

MACOSA also explored the feasibility of assessing social studies achievement.

[8] See pp. 6 - 7 of this report for exact titles and bibliographic information.

Feasibility Study on Assessing Student Achievement in Social Studies

Purpose The purpose of the study was to determine the feasibility of assessing social studies skills and knowledge among Alberta students.

Procedures The researchers reviewed the literature on evaluation in social studies, examined available tests, identified objectives of the 1971 Alberta social studies program* and determined the degree of specificity of these objectives. To provide additional data they interviewed knowledgeable social studies educators and surveyed a sample of social studies teachers by means of questionnaires.

Findings The literature indicated that it is possible to assess achievement of the knowledge and skills objectives of social studies programs, provided that the objectives are stated quite specifically and the curriculum assigns skills objectives to specific grade levels. According to the research reviewed, assessing achievement of affective objectives (attitudes and values) is more difficult. Even so, the researchers reported that, because the 1971 Alberta social studies curriculum particularly emphasizes the valuing process, assessment in the affective domain is especially important to Alberta's testing programs.

The researchers found that available standardized tests developed for American schools could not be adapted to fill Alberta's needs because the Alberta social studies program is in many respects unique in North America. They identified a few tests that could be used to assess knowledge and skills, including the high school social studies achievement test developed by the Student Evaluation and Data Processing Services Branch of Alberta Education (SEDPS).

Educators responding to the researchers' survey reported that expertise for developing test items related to knowledge and skills objectives was readily available. Uncertainty prevailed regarding development of tests pertaining to attitudes and values, but there was some hope that such tests could be developed if a group of dedicated and interested teachers were set to work on the task.

The educators tended to agree that it would be feasible to conduct an Alberta assessment of student achievement in the skills area. There was less agreement on whether knowledge objectives should be assessed on a province-wide scale since the curriculum guidelines permit a considerable amount of local decision making about what factual material is to be included in social studies programs. It was noted that the curriculum was currently being revised with the goal of providing more clearly specified content for each grade level and more emphasis on Canadian content.

Conclusions and Observations The study concluded that assessment of social studies achievement is hindered primarily by the lack of specificity in the 1971 social studies program. Few available tests were judged suitable for use in Alberta, but there were indications that some new tests could be developed.

* At this time (1977) the Curriculum Branch of Alberta Education was extensively revising the social studies program. The new program was to be available to schools in 1979.

The researchers proposed two alternatives — a partial assessment or a more comprehensive assessment. The first plan involved an assessment based on the 1971 curricular objectives at the grades 4, 7, 10 and 12 levels. New tests were to be developed for grades 4 and 7 but the current SEDPS high school achievement test was to be used for grades 10 and 12.

The second plan, which would be carried out after the revised social studies curriculum became available, involved the development and validation of tests to assess the more specific knowledge, skill and affective objectives which were to be stated in the new curriculum.

Action by MACOSA MACOSA opted for a modified version of the first plan; that is, using only the SEDPS high school achievement test for grades 10 and 12 and not assessing achievement at the elementary level. A second outcome of this preliminary study was a project to develop models of tests for the affective domain of the social studies.[10] The achievement study is described in Chapter 5 and the affective test development study is described in Chapter 6.

Science

The feasibility of assessing science achievement was also looked at by MACOSA.

Feasibility Study on Assessing Student Achievement in Science

Purpose The purpose of the study was to determine the feasibility of assessing the scientific knowledge and skills of Alberta students.

Procedures To develop a plan for assessing the science achievement levels of Alberta students the researcher reviewed the literature on evaluation of students' scientific knowledge and skills, identified the objectives of the Alberta science program, examined available science tests to determine their suitability for use in Alberta, and finally interviewed knowledgeable science teachers and their supervisors regarding the proposed assessment program.

Findings and Conclusions The literature clearly indicated that the knowledge and skills objectives of the Alberta science program could be tested. The researcher found that no suitable tests were available for the elementary level because the curricular objectives of the Alberta elementary science program are uniquely organized in a spiral form, but that a published standardized test would be suitable for assessing science achievement in the junior high school because that program in Alberta is similar to many other junior high school science programs throughout North America. By using an existing test, MACOSA could avoid duplication of effort and take advantage of established norms. For the grade 12 level where a number of tests of specific areas of science (biology, chemistry, physics) had already been developed, the researcher advised that a general science achievement test that would provide data on the overall level of student achievement in science would be most appropriate.

[10] The exact titles of these studies appear on p. 7.

The educators who were interviewed thought it would be desirable to conduct a science assessment. They agreed that the tests should provide a balance between knowledge and skill items, but pointed out that there would probably be disagreement about attitude items.

Recommendations The study recommended that:

1. Science tests should be administered in grades 3, 6, 9 and 12.
2. For grades 3, 6, and 9 the test items should reflect the program objectives.
3. A general science achievement test should be administered in grade 12.

Action by MACOSA The MACOSA science achievement study, which is described in Chapter 5, closely followed the procedures recommended by this preliminary study.

CHAPTER 4

Studies on Factors Affecting Student Achievement

The Alberta *Goals of Basic Education* statement (1978) identifies two types of goals — goals of schooling and goals of education — and states that school personnel should take a primary responsibility for achieving the goals of schooling and a shared responsibility for achieving the goals of education. By making a distinction between "schooling" and "education", the Goals document acknowledges that both educational and environmental factors have a considerable effect on student achievement and that the influence of the school is modified by other agencies and non-school elements.

MACOSA also recognized that educational and environmental factors have a significant effect on student achievement and took these factors into account when making judgements about the adequacy of student performance levels in the various subject areas. However, MACOSA could not judge exactly what effects these factors have on achievement. Factors related to student achievement are difficult to define and measure, vary from school to school or over time, and their effects may be ambiguous. Furthermore, several of these characteristics can and often do exist simultaneously. For example, low achievement scores could be attributed to all or several of the following factors: socio-economic status, which varies from one area to another; teaching methods, which vary a great deal and are difficult to measure; and cultural attitudes towards schooling, which may vary over time. Because of these complexities, it was not feasible to gather information about such factors along with the test results, and therefore MACOSA did not have available conclusive data which could explain students' strengths and weaknesses.

MACOSA commissioned two reviews of the literature (Fedigan, 1978; Gay, 1978), both of which suggested that environmental or non-school-based elements (which are often beyond the control of educators) have a substantial relationship to achievement while educational or school-based elements may have a less substantial and more inconsistent relationship. Some researchers (e.g., Coleman) have even concluded that teachers do not make a difference, and that classroom instruction has no significant effect on the outcome of schooling. This theory is further reinforced by findings that student aptitude accounts for 15 to 50 percent and home environment accounts for up to 64 percent of the variance in learning. Although most educators would question such an extreme position, these findings can help decision makers to gain a more realistic perspective from which to view student achievement.

All studies reported in MACOSA's reviews of the literature have examined some aspect of the relationship between or among educational outcomes and environmental and educational factors. "Environmental" factors are those background influences which students bring to school such as type of community including the mix

according to social, economic, racial, educational and achievement levels; attributes of the home; and qualities of the student body according to its academic readiness, motivation, and social behavior. "Educational" factors include such factors as teacher characteristics, the structure of the teaching-learning situation, school facilities and school personnel services.

Both reviews point out the significance of the interaction between educational and environmental factors. In practical terms, knowledge about such interaction could permit educators to maximize learning by matching instructional techniques with student characteristics. (See Figure 1.)

Figure 1. Interaction Between Educational and Environmental Elements

The studies reviewed found that many factors were related to student achievement. It is important to note, however, that these factors may or may not actually cause the results reported. If student achievement increases consistently as attendance increases, there is a positive relationship and if student achievement increases consistently as drug usage decreases, there is a negative relationship. This does not mean, though, that good attendance necessarily improves student achievement or that the use of drugs is detrimental to student achievement. Cause and effect cannot be inferred from every such relationship.

Some studies have shown that achievement in the knowledge and skills areas is positively related to median family income (0.88),[11] living with both parents (0.83), use of books in the home, teachers' salaries, library facilities in the school,

teacher satisfaction with students, teacher expectations, and teacher control of classes; and that such achievement is negatively related to housing density (-0.69), family size (-0.55), juvenile delinquency, and lack of library services (-0.53). Attendance at school, amount of money spent per pupil and class size have been shown to be ambiguously related or not related to achievement in the knowledge and skills areas. Studies have also found that students' attitudes are positively related to parents' attitudes towards education and teacher satisfaction with students, and negatively related to drug usage, part-time work and class size.[12]

Research on school-based elements (teaching behavior, classroom climate, and student characteristics) has been relatively unproductive to date, but the scope of such research has been quite limited. Studies done between the mid-1940's and early 1960's found that teacher traits such as enthusiasm and warmth appeared to promote effective learning. During the 1960's most researchers were committed to the indirect teaching method — that is, making use of student contributions and being democratic rather than authoritarian — and their studies attempted, often without success, to show that this approach was effective in terms of student achievement. More recently, several reviews of the literature have also indicated that indirect teaching is less effective than other teaching approaches. However, five promising teaching behavior variables have been supported by significant results in several studies. They are: clarity, variability, enthusiasm, task orientation or businesslike manner, and student opportunity to learn what is being tested. A further study found consistent results for only four variables. Allowing students the opportunity to learn what is being tested, task orientation, and direct questioning had positive effects on achievement, and teacher criticism had a negative effect.

The Texas Teacher Effectiveness Study (1973-1976) found that the most consistently effective teachers tended to use traditional (or direct) teaching styles. This study also found that students of low socio-economic status seemed to learn best in a controlled atmosphere, while students of higher socio-economic status could also learn well in a less structured environment where teachers challenged and criticized the students more frequently. Many other studies have reinforced the idea that direct teaching (teacher control, emphasis on task, seatwork) is generally effective in lower elementary grades, where basic skills are being acquired, and with students of low socio-economic status. In other words, it seems important that teaching behaviors be matched with student characteristics for more effective learning.

Other studies have explored the differences in methods used in different school subjects, particularly English and mathematics. Different teaching approaches appear to be more effective for certain subjects — for example, a direct approach

[11] When correlation coefficients describing strengths of relationships between specified factors and achievement are known, they are indicated in brackets.

[12] It should be noted that only one study was reviewed on the relationship between class size and students' attitudes.

may be more suitable in a mathematics class — but the results are not as conclusive as those for the other studies described here.

In general, the research showed that direct teaching is quite effective in certain situations, but that indirect teaching could be more successful with older, higher socio-ecomomic status, or higher ability students, and in subject areas like English. It is not yet known whether these findings are readily applicable to schools in general, and questions remain about the effects direct teaching may have on outcomes other than achievement such as attitude, creativity or social skills.

In 1968 Rosenthal and Jacobson published their controversial "Pygmalion" findings; that is, that teacher expectations strongly influence student achievement. Rosenthal claims not only that the Pygmalion effect exists, but also that it applies to teacher expectations in relation to many student characteristics including sex, ethnic background and socio-economic status as well as aptitude. Teachers' different expectations of boys and girls serve to perpetuate cultural sex differences; and the cycle of low socio-economic status, low ability, low teacher expectations and low achievement continues to maintain social stratification. Such large-scale intervention programs as Headstart in the United States have attempted to break this chain of events by altering teacher expectations and they have succeeded to a certain degree. (Interestingly, in such programs a direct, highly structured teaching approach has again been found to be the most successful.)

Studies that have attempted to find out what curriculum is most effective in terms of achievement or outcome have been generally inconclusive. New programs or experimental curriculum packages, including programmed or television instruction, have shown no consistent pattern of achievement gains over traditional materials. The use of individualized instruction, both as a curriculum and a teaching approach, seems to be more related to a commitment to the recognizing of individual differences than to efficiency in producing high student achievement levels.

The mastery learning or criterion-referenced approach to individualization, in which each child is given as much time and support as necessary to learn a certain skill or body of information, has been found to be the most efficient and productive approach in terms of use of time. In this case "time" refers to time spent in learning, which is not necessarily the same as time spent at school. Using the mastery learning approach, most students succeed and become more efficient learners as well. They learn *how* to learn. As a result individual differences diminish during the course of schooling, whereas in traditional classrooms individual differences increase over time. These findings about mastery learning reinforce the findings of other studies in which such factors as attention to task and clarity of directions are found to have a positive relationship to achievement.

Consideration of the factors affecting student achievement gives rise to some

interesting problems. Most research studies today are directed towards the single pragmatic goal of efficiency; yet educators are probably not willing to use extreme teaching methods to achieve high academic performance levels. Similarly, if the research were to show that knowledge and attitude objectives are independent or unrelated to each other and require opposing teaching styles, educators would have to decide which objectives they wish to emphasize. Because schools today involve high financial investments, many people favor a business-and-industry approach in which "accountability" is the watchword. At the other extreme, some people argue that teaching is an art form which cannot be analyzed scientifically.

If we accept Coleman's statement that non-school-based influences are much more significant for learning than are school-based factors, the implications for schools are both negative and positive. Economic and political decisions made on the basis of this belief would greatly decrease the role of the school in our society, but at the same time the school would be freed from its present burden of trying to be the panacea for social ills ranging all the way from malnutrition to pollution.

In summary, the two reviews of literature commissioned by MACOSA found that the small amount of research which has been done to date tends to be inconclusive in terms of any consistent effect of environmental or educational factors on student achievement. Nevertheless, readers are advised to review the reports of the following achievement studies (Chaper 5) with such factors in mind.

CHAPTER 5

Achievement Studies
Levels and Adequacy of Student Performance

This chapter summarizes the purposes, procedures and findings of each MACOSA achievement study and states the Advisory Committee's conclusions and observations about those findings.[13]

Alberta Grade 3 Achievement Study

Purposes The purposes of the study were:

1. To establish 1977 benchmarks for grade 3 achievement in reading, language and arithmetic.
2. To make inferential comparisons, where possible and tenable, between 1977 achievement and the norms of the mid-1950's. To make these comparisons the researchers also used the results of a companion study, the *Edmonton Grade 3 Achievement Study* (Clarke and others, 1977).

Procedures The researchers used data from the following commercial standardized tests, gathered first in 1956 (the *Edmonton Grade 3 Achievement Study*) and again, using the latest versions of the tests, in 1977: Gates MacGinitie Reading-Vocabulary and Comprehension, and California Achievement Tests of Reading, Mathematics and Language. For the 1977 study a number of teachers and university educators assessed both sets of tests to determine their applicability to the present Alberta curriculum, and before processing the results the researchers eliminated items which were judged invalid. Few items were deleted in the language arts tests, but about a dozen were removed from the mathematics test.

The tests were administered during May and early June of 1977 to all grade 3 students in a 10 percent random sample of Alberta schools offering instruction at the grade 3 level.

The study compared the average scores of various groups in the following ways: 1977 achievement scores for the four large urban districts in Alberta compared with scores of rural counties, divisions, and districts; Edmonton Public School District 1977 norms compared with 1956 norms in the same district, with 1977 averages in other large urban districts, and with 1977 averages in all other school systems in Alberta; and finally all of the scores compared with norms reported by test publishers.

Findings Scores were similar on tests administered in 1977 in four large urban districts, including Edmonton Public. The average scores of this metropolitan group were generally equal to or slightly below those of the non-urban groups in counties, divisions and rural districts. (This analysis excluded small city districts because this group was not adequately represented in the sample.)

The study made twelve comparisons. Five comparisons showed a difference which

[13] Appendix C provides information about distribution and sources of the various MACOSA studies.

favored non-urban students and one favored students in the metropolitan group. Edmonton Public School District scores for 1977 were similar to those from the same district in 1956. There were slight gains in reading and language test scores and a very small decline (0.42 raw score points) in arithmetic. Scores on the California Test of Mental Maturity averaged over four points higher in 1977 than in 1956. All scores exceeded the publishers' norms, which were based on samples from United States student populations between 1942 and 1970. The Edmonton Public District group exceeded the publishers' norms in both 1956 and 1977, and other Alberta groups did so in 1977.

Conclusions and Observations Regarding achievement in grade 3, differences between rural and large urban groups which may have existed previously appear to have been eliminated. Conventional wisdom has held that achievement in large city schools in Alberta is superior to achievement in non-urban schools, and some limited evidence has supported this claim. A study by Carmichael and Rees (1955) showed that during the mid-1950's reading scores for grade 4 students in city and town schools were significantly higher than scores in graded and ungraded rural schools. If differences favoring the urban student also existed in language and arithmetic performances, it is evident that improvements have resulted. The gains reported in this study, which are largely due to improvements in rural areas, may indicate that greater equality of educational opportunity has been achieved.

The *Edmonton Grade 3 Achievement Study's* assessment of changes over a 21-year period in the Edmonton Public School District indicated that performance levels in reading, language and scholastic aptitude have been maintained or have improved despite the fact that the 1977 group was, on the average, 1.72 months younger. Achievement in arithmetic has remained essentially the same. Also, the 1977 achievement levels of Alberta grade 3 students in the areas measured by the tests were higher than the achievement of the United States comparison groups in previous years.

If provincial data from the mid-1950's had been available in the form used in the *Edmonton Grade 3 Achievement Study*, direct comparisons could have been made. In the absence of such information, MACOSA could infer only that achievement levels of the Alberta grade 3 student population have generally improved.

Alberta Language Arts (Reading and Writing) Achievement Study

Purposes The purposes of this study, entitled the *Alberta Language Arts (Reading and Writing) Achievement Study*, were:

1. To provide information about current levels of student achievement in reading and writing for selected objectives at the grades 3, 6, 9, and 12 levels in Alberta schools.
2. To provide a data base for future assessments.

Procedures At the time this study was being conducted, the Alberta Education curriculum guides for language arts stated the program objectives in rather general terms. The researcher, therefore, began by compiling a comprehensive list of specific objectives suitable for use in developing tests of reading and writing at the grades 3, 6, 9 and 12 levels. The steering committee for the study, with the assistance of the Curriculum Branch of Alberta Education and its language arts coordinating committee, then ranked the measurable objectives in terms of their importance to the total program. When the committees had established priorities, the researcher had available a selected list of objectives for which appropriate test items could be developed.

The researcher used standard piloting and field testing procedures to investigate administrative and scoring techniques and to determine what revisions were needed. The manual scoring required by the writing tests presented a special challenge. To ensure fairness and consistency, two specially trained markers scored each composition according to two different marking scales. In the event of a major discrepancy between the two scores, a third marker scored the composition.

MACOSA considers its reading and writing achievement tests to be superior to those available commercially because the MACOSA tests are more valid for the Alberta curriculum and the writing test employs a unique marking technique which ensures a high degree of fairness and impartiality in the scoring of students' writing.

The reading and writing tests were administered at each of grades 3, 6, 9, and 12 in a random sample of 284 schools on May 17, 1978.

Findings — Reading Table 1 shows the levels of performance in reading. Not all selected skills were tested in each grade, either because they were inappropiate for that grade level or because of the necessity of limiting the length of the test. Therefore, blanks appear in some columns of the table.

At some or all of the grade levels students were asked to determine the meanings of new words by applying knowledge of sounds (phonics) and word parts (structure), and by referring to the sense of the sentence (context). Grades 3 and 6 students were also tested on their knowledge of the meaning of words they had learned to read by sight.

A second level of the test required that students demonstrate their understanding of details, relationships and main ideas which are directly stated in connected writing (sentences, paragraphs or longer passages) and in graphic materials such as graphs, maps, and charts. A further division of this level required that students identify implied details, relationships and main ideas in selected reading materials. At some grade levels students were also asked to grasp implied purposes or tones and to use reading material to draw conclusions, make predictions, and form judgements.

Table 1

Reading Achievement Levels: Percentage of Correct Responses

	Grade			
	3	6	9	12
Word Level				
Meaning through analysis of:				
Sounds (phonics)	81	—	—	—
Word parts (structure)	78	47	67	62
Sense of sentence (context)	77	76	75	70
Knowledge of meanings of sight words	68	68	—	—
Connected Writing Level				
Comprehension of clear and direct statement of:				
Details	63	72	56	—
Relationships	66	61	63	70
Main Idea	68	72	68	76
Use of reading material to grasp implied:				
Details	68	57	—	73
Relationships	78	67	66	57
Main Idea	67	57	60	75
Purpose	—	—	62	71
Tone	—	—	63	63
Use of reading material to draw conclusions, make predictions, and form judgements	—	59	68	73

The lowest average score on a cluster of reading objectives at any grade level was 47 percent and the highest average score was 81 percent.[14] Grade 3 average scores all closely approached or fell within the upper half of this range. Highest achievement was recorded in obtaining word meaning through analysis of sounds (phonics) and lowest achievement was in comprehension of clear and direct statements of details.

At the grade 6 level average scores for five of the ten skills tested fell within the total range described above. The highest score for grade 6 was for obtaining the meaning of a word from context and the lowest score was for obtaining the meaning of a new word through analysis of its parts (roots, prefixes and suffixes). Relatively low achievements levels were evident in identifying implied details and main ideas and in using reading materials to draw conclusions, make predictions and form judgements.

Grade 9 scores fell within or near the upper half of the total range on eight of the ten skills tested. The lowest scores occurred in identifying implied main ideas and comprehending clues and direct statements of details. As in grade 6, the highest score was achieved in obtaining the meaning of a word from context.

Grade 12 scores approached or fell within the upper half of the total range on all but one of the ten skills tested. The lowest achievement was recorded for identifying implied relationships. The highest score was on comprehension of clear and direct statements of main ideas, and this was closely followed by achievement in identifying implied main ideas.

Conclusions and Observation - Reading MACOSA judged that performance in reading was generally satisfactory for all four grade levels. The grade 3 performance in phonics, which was judged to be particularly satisfactory, undoubtedly reflects an emphasis on this skill in primary reading programs. In several instances there is evidence of improvement or at least of maintenance of skills as students progress through the grades. For example, levels of achievement in using reading materials to draw conclusions, make predictions and form judgements increase from grade 6 to grade 12. Similarly, performance levels in using the content to obtain meanings of words are stable across all four grades. Performance in comprehension of clear and direct statements of main ideas fluctuates slightly from grades 3 to 9 and increases in grade 12. These performances are also considered satisfactory.

[14] Although valid comparisons across grades cannot be made in the mathematics and science studies, such comparisons are quite legitimate and by design in the reading and writing studies.

The somewhat surprising weakness of grade 6 students is using word parts to obtain meaning is judged to be unsatisfactory, as is the low grade 9 performance in identifying directly stated details. Regarding the grade 9 performance on directly stated details, MACOSA noted that the test questions related primarily to maps and charts. Although reading maps and charts may be more difficult than reading prose, an ability to read maps and charts is considered essential for that grade level. Weaknesses in identifying implied elements of a passage were apparent in each of grades 6, 9, and 12. MACOSA recognized that it is more difficult to draw out implications, but nevertheless considered performances in this area unsatisfactory for the following skills and levels: drawing conclusions, making predictions and forming judgements in grade 6; identifying implied main ideas in grade 9; and identifying implied relationships in grade 12.

Findings - Writing The MACOSA writing tests assessed two kinds of skills. A short answer section asked students to find and correct errors in capitalization, punctuation and spelling, and to write sentences. A second section required that each student write a composition, the nature of which varied according to the grade level being tested. Students were asked to write a story in grade 3, a story or description in grade 6, an exposition in grade 9, and an argument in grade 12. The markers scored each type of writing with different characteristics in mind regarding overall impact.

To score the longer writing assignments, the markers used two different scales called "descriptive" and "holistic". The descriptive scale used four-point descriptors for each of ten selected objectives. For example, for the objective of vocabulary breadth a score of "4" indicated vivid and descriptive word choice while a score of "3" indicated ordinary but functional word choice. The holistic scale indicated an overall impression of the writing. Positive impressions were described as impressive (a score of 5), commendable (4), and functional (3), while negative impressions received a score of 1 or 2. The functional category, or the "3" rating, included performances which might be described as average or low average. Students who received the functional rating were able to communicate on a basic level but needed instruction in some areas.

Tables 2 and 3 show the levels of student achievement in writing. Paragraphing was not assessed in grade 3 because this skill is not emphasized until later in the program. In the interests of brevity, capitalization was omitted from the short answer sections of the grades 6, 9, and 12 tests.

> For all four grades, scores were low on the short answer section of the test, which assessed abilities in capitalization, spelling, punctuation and sentence structure. On the longer writing assignment, performances varied from grade to grade but, on the basis of the descriptive scoring scale,

Table 2

Writing Achievement Levels: Percentage of Correct Responses

Skill	Grade 3	6	9	12
Short Answer Assignment				
Finding and correcting errors in:				
Capital letters	70	—	—	—
Spelling	47	44	55	48
Punctuation	40	50	36	58
Writing sentences	66	28	24	18
Longer Writing Assignment — Descriptive Scoring				
Type of Writing	Story	Description or Story	Exposition	Argument
Limitation of content to topic	94	93	95	90
Development of topic in terms of directions	80	86	72	47
Sentence structure, capitalization and punctuation	51	40	61	50
Paragraph (coherence)	—	58	56	55
Paragraph as a unified part of the whole work	—	30	49	66
Word use in terms of its function	85	88	71	61
Breadth of vocabulary	84	88	86	60
Spelling	68	70	84	74
Total effect or impression	73	63	69	73
Legibility	88	74	89	78
Averages for Both Assignments	70	60	60	56

Table 3

Longer Writing Assignment — Holistic Scoring: Percentage of Students Assigned Various Ratings

	Grade			
	3	6	9	12
Rating on 5-point Scale				
Impressive (5)	2	5	7	3
Commendable (4)	23	22	45	22
Functional (3)	58	65	41	49
In need of remediation (2)	15	6	6	22
Insufficient material (1)	2	2	1	4
Aggregations of Above Ratings (%)				
Commendable or better (Scores of 4 or 5)	25	27	52	25
Functional or better (Scores of 3, 4 and 5)	83	92	93	74

students in all grades achieved high scores in limiting content to the topic and writing legibly.

On the longer writing assignment, grade 3 students also received relatively high scores for developing the topic in terms of directions, using words correctly in terms of sentence functions, and using vivid, descriptive words, while students received low scores for capitalization, punctuation and spelling. Combined results for grade 3 on the short answer section and the descriptive scoring of the longer writing assignment produced an average score of 70 percent. Application of the holistic marking scale resulted in ratings of functional or better (scores of 3, 4, or 5) on 83 percent of the stories. Twenty-five percent of the stories were deemed commendable or impressive (scores of 4 or 5).

Grade 6 students had high scores on developing the topic in terms of directions, using words appropriately and correctly in terms of audience, and using vivid, descriptive words and phrases, as well as on limiting content to the topic, and legibility. Grade 6 students received low scores on writing complete sentences, using capital letters, punctuating, and paragraphing. Combined results for the two sections of the test produced a grade

average of 60 percent, which is considerably lower than the grade 3 average but about the same as the averages for grades 9 and 12. In the holistic scoring of the grade 6 writing assignments, 92 percent of the compositions were rated as functional or better, and 27 percent were rated as commendable or impressive.

Grade 9 students achieved high scores for limiting content to the topic, legibility, spelling, and breadth of vocabulary. They had low scores on sentence structure, capitalization, punctuation, and paragraphing. Combined scores for the total test produced an average of 60 percent. Holistic scoring of the longer writing assignments resulted in ratings of functional or better on 93 percent of the expositions, with ratings of commendable or impressive on 52 percent. The number of compositions given ratings of 4 or 5 was much greater in grade 9 than at any other level.

Grade 12 students received high scores for spelling and achieving effect through a consistent point of view, as well as for limiting content to the topic and legibility. Performances were lower on sentence structure, capitalization, punctuation, paragraphing, word use, breadth of vocabulary, and developing the topic in terms of directions. The combined results for the two sections of the test produced a total test average of 56 percent. In the holistic scoring of the grade 12 writing assignments, 74 percent of the compositions were rated as functional or better, as compared to 83, 92 and 93 percent in this category in grades 3, 6, and 9 respectively. Commendable or impressive ratings were given to 25 percent of the grade 12 essays.

Conclusions and Observations - Writing MACOSA judged writing at the grade 3 level to be satisfactory. Since the curriculum emphasizes reading more than writing in the first three grades, MACOSA concluded that writing performances should be judged in terms of this lesser emphasis. In other words, weaknesses in capitalization, punctuation and sentence writing in grade 3 might be regarded as a direction for skill development in future grades.

Achievement levels in the two sections of the grade 3 writing test presented a marked contrast, which suggested that students are more successful at using their own language for their own purposes than at manipulating structured assignments. MACOSA speculated that the current curriculum, which emphasizes the practical use of language more than the study of the forms of language (grammar), may be having some impact.

In view of the fact that the grade 6 curriculum for language arts equally stresses reading, writing, speaking and listening, MACOSA judged the writing achievement levels for that grade to be marginally satisfactory. There was a sharp contrast between grade 6 achievement levels on the two sections of the writing test. When writing words of their own choosing, as on the longer writing assignment, grade 6

students presented themselves as better spellers than when asked to identify misspelled words chosen by the test developers. However, results of both sections of the test indicated weaknesses in writing sentences, using capital letters and punctuating. The grade 6 results ranged from a low of 28 percent to a high of 93 percent. This broad range suggests that certain elements of the writing process may be receiving more attention than others. For example, the average score of 30 percent for paragraphing might be attributed to a lesser emphasis on this skill at the grade 6 level. At the same time, though, it should be noted that organization of thought into compact units is not an easy task at any level.

MACOSA judged writing at the grade 9 level to be satisfactory. Regarding the low average score on sentence writing in the short answer section of the test, MACOSA noted that students were given a rather difficult assignment which involved writing mainly complex and compound-complex sentences.

MACOSA judged grade 12 writing to be unsatisfactory. Although it was recognized that every high school graduate may not be capable of producing completely error-free and polished writing, members of MACOSA agreed that more than 74 percent (see Table 3) of the students should have performed at the functional[15] level or better.

The following points summarize MACOSA's judgements on the writing test results taken as a whole.

1. Alberta students are generally more successful at using their own language in writing than at finding errors in or manipulating structured assignments.
2. On the longer writing assignment performances were generally strong at all levels for limiting content to the topic, spelling, total effect or impression, and legibility; and at all levels except grade 12 for developing the topic in terms of directions, word use in terms of function, and vocabulary.
3. Achievement was low on capitalization, punctuation and sentence structure in both sections of the test and for all grade levels. On the longer writing assignment the duress of the test situation may not have permitted much time for proofreading, correcting and re-organizing. Nevertheless, scores in these areas indicate some cause for concern.
4. The scores for paragraphing (not tested in grade 3) increased progressively from grades 6 through 12. MACOSA speculated that mastery of the rather complex skill of paragraphing may be highly dependent upon reaching a certain level of maturity.

[15] Students who received the "functional" rating were able to communicate on a basic level but needed instruction in some areas.

Alberta Mathematics Achievement Study

Purposes The purposes of this study, entitled the *Alberta Assessment of School Mathematics*, were:

1. To provide information about current levels of achievement in mathematics skills and knowledge among students in Alberta schools.
2. To provide a data base for future assessments.

Procedures The researcher began by preparing a table of specifications which prescribed for each grade level the number and types of text exercises required to assess achievement in various content areas of mathematics and at three thought levels — knowledge, comprehension and application. Next, teachers reviewed 1500 items from similar assessments done in Canada, the United States and Britain and evaluated the items in terms of clarity, fairness and representativeness. Other teachers and the steering committee for the study reviewed the items again after they were revised. These activities, supplemented by information from pretesting with small groups, provided the basis for the final selection of items.

Students at each grade level responded to 120 questions. About 70 percent of the questions were multiple choice; the remainder were open-ended. Between 54 and 92 of the 120 questions were appropriate to the target grade, and the others were review items appropriate to lower grades or preview questions suitable for higher grades. Review and preview items were administered as common items in several grades to provide cross-sectional information.

On May 17, 1978, the tests were administered to a random sample of schools consisting of 102 schools offering grade 3; 101 schools offering grade 6; 56 schools offering grade 9; and 50 schools offering grade 12. A form of matrix sampling was employed so that each student did not respond to all 120 test items; instead, different sets of test exercises were randomly assigned to each student.

The study recorded percentages of correct responses for the total groups and for sub-groups defined according to sex, school size, and, for grade 12, type of high school mathematics.

Findings Tables 4 and 5 present summary information about current mathematics achievement levels in grades 3, 6, 9, and 12 in Alberta. The reader is cautioned not to draw conclusions or make comparisons between grades using the data displayed in these tables. The results from the four grades are combined on the tables for the sake of brevity and convenience. Since common or equivalent test item results are not tabulated across grades, it is inappropriate to make such comparisons. It may also be misleading to use the absolute values of the percentages reported as the only

criteria in judging the quality of achievement. The groups who made judgements about achievement levels applied other criteria, such as the perceived difficulty of the concepts, the curricular emphases or appropriateness of the concepts, and the nature of the test items.

Blanks appear in the consumer mathematics section of Table 4 because that subject area was not tested in grades 3, 6, and 9. The total number of items for each grade as indicated in both tables refers to items directed towards that level. The rest of the items written at that grade level were preview or review questions, as described above.

Table 4 indicates that achievement in the number strand is relatively high. For grades 3 and 6 the test items in this category were predominately on knowledge of number and computations such as addition, subtraction, multiplication and division. In contrast, achievement levels in geometry, measurement and consumer mathematics were consistently lower.

Table 4

Average Mathematics Achievement Levels
For Content Strands:
Percentage of Correct Responses

Content Strand	Grade 3	6	9	12
Number	70 (49)*	73 (25)	70 (15)	69 (8)
Algebra	68 (9)	63 (7)	50 (21)	53 (27)
Geometry	54 (20)	55 (14)	66 (11)	53 (10)
Measurement	64 (10)	57 (6)	65 (5)	60 (7)
Statistics	58 (4)	48 (3)	79 (3)	53 (3)
Consumer Mathematics	— (0)	— (0)	— (0)	51 (18)
Total	65 (92)	64 (55)	62 (55)	53 (73)

* The number of items in each category appears in parentheses.

Table 5 shows the test results by thinking processes, ranging from simple to complex. The thought levels are knowledge (simple recall), comprehension (understanding), and application (organizing and reorganizing information as in solving problems). Achievement at the application level is relatively low for each grade, an indication of students' general difficulty with problem solving tasks. Achievement is higher on less demanding tasks involving either remembering or understanding.

Table 5

Average Mathematics Achievement by Thought Level: Percentage of Correct Responses

Thought Level	Grade 3	6		12
Knowledge	73 (36)*	72 (24)	64 (19)	61 (14)
Comprehension	61 (41)	58 (17)	66 (21)	55 (21)
Application	57 (15)	57 (14)	53 (15)	53 (38)
Total	65 (92)	64 (55)	62 (55)	55 (73)

* The number of items in each category appears in parentheses.

Detailed results of the mathematics study in terms of content areas may be summarized as follows.

> *Number.* The results on basic number operations (adding, subtracting, multiplying and dividing) varied greatly, ranging from a low of 11 percent on some items to a high of 99 percent.

Grade 3 students had high achievement levels on the fundamental skills of primary school mathematics. Students' scores ranged from 91 to 98 percent on items dealing with the cardinal number of a set, reading three-digit numbers, counting by tens, ordering four-digit numbers, and the basic facts of addition and subtraction. Performances were much lower on exercises dealing with common and decimal fractions, ranging from 11 to 72 percent. At the low end were decimal tasks like asking that 0.3 be read as three-tenths, and at the high end were tasks on recognizing common fractions.

Grade 6 students continued the strong start made on fundamental skills in earlier grades, obtaining scores in the 90's on exercises dealing with basic facts of multiplication and division, adding and subtracting and multiplying multi-digit numbers, reading five-digit numbers, and rounding to the nearest ten. As in grade 3, the grade 6 scores were low on common fractions and decimal fractions and also on integers.

Areas which showed high performances in grades 3 and 6 remained strong in grade 9 (particularly whole number work). Grade 9 performances on place value typically ranged from 85 to 95 percent and performances on whole number application were in the 80's. Results on percentage and integers were somewhat mediocre. As at the grades 3 and 6 levels, the results on common and decimal fractions failed to reach the level of whole number work, although computing with decimal fractions rose significantly.

For grade 12 the picture in number work was about the same as for grade 9; areas which were low in grade 9 did not improve significantly in grade 12.

Algebra Although the algebra program receives little emphasis before grade 9, students in grades 3 and 6 showed some understanding of fundamental algebraic concepts. Over 80 percent of the grade 3 students tested showed an understanding of the use of the box "□" in a simple algebraic sentence while two-thirds of them could translate an English statement into a sentence containing a box. It is encouraging that one-third of the students indicated an understanding of the concept "more than" and the accompanying symbolism, "$>$". By grade 6, two-thirds of the students were using this idea on inequality correctly, while grades 9 and 12 showed 90 percent correct usage.

In spite of the fact that the grade 6 algebra program focuses on word problems, the best performance for a single item on problems containing three or four sentences and one or two operations was about 75 percent. Grade 9 students performed slightly better on these problems. Over half of the grade 6 students could solve a simple equation involving "x", but grades 3, 6, and 9 all performed at a low level in solving an equation when x was negative.

Grade 9 results were low on linear expressions like $r + s - (r - s) = \square$, and where the answer required the use of a variable. On standard "age" problems grade 9 students scored higher but any variation which made the problem slightly unusual brought about marked drops in the scores. Only about half of the grade 9 students understood the linear function; that is, for example, finding the value of x given a value of y in an equation like $5x + 6y = 24$.

Grade 12 students had high scores on simple algebra involving equality, inequalities, simple verbal problems, evaluating polynomials, and simultaneous linear equations. About 75 percent could interpret a formula involving a linear function. About half of the grade 12 students could deal with quadratics, graphs of linear equations, exponents, division of polynomials, and solving rational expressions. In several other areas grade 12 students performed at slightly less than 50 percent or even much lower; for example, in conics, composite functions, exponential functions, and trigonometry. Relatively low scores characterized the grade 12 results both on problems where answers were in terms of variables and on problems which took a slightly unusual form.

Geometry Geometry, together with number and algebra, is a fundamental part of any mathematics program from the early grades on. Performance levels varied, probably in relation to the amount of emphasis each area received in the school program. For example, one exercise on points, lines and space in two dimensions had results of 80, 83 and 88 percent in grades

6, 9 and 12 respectively, while a similar exercise in three dimensions had results of 34 and 36 percent in grades 9 and 12.

In general, performance levels were relatively high on angles, triangles, circles and congruence. In contrast, the following topics were not handled nearly as well: similarity, line relationship (parallelism and perpendicularity), polygons, transformations and symmetry. In a number of exercise irrelevant perceptual cues seemed to mislead the students. For example, a significant percentage of students did not recognize a square supported on one of its corners. Students also experienced difficulty with word problems, which may have required a high level of reading ability.

Measurement Most of the measurement exercises tested familiarity with the metric system or the application of metric units in geometric settings. Telling time, figuring with money, and measuring length had strong beginnings in grade 3 and rose to 90 percent performance in grade 6. Questions on perimeter, area and volume fared less well, resulting in more average performances, even at the grade 12 level (70 percent). In particular, performances in grades 9 and 12 on exercises dealing with inter-relationships among length, area and volume resulted in quite low scores (30 to 40 percent). As well, over one-quarter of the grade 12 students apparently are not familiar with the inter-relationships among metric prefixes.

Statistics All items in statistics dealt with comprehension or application of knowledge. Grade 3 students had high scores on picture graphs. Only one-third of the grade 3 students were able to construct a bar graph. Grade 6 students had high scores on bar and line graphs and scored about 50 percent on the circle graph. Half of the grade 6 students were able to solve a problem based on a multiple line graph. Grade 9 performances followed the trend, with almost 80 percent correctly interpreting a circle graph. Problems based on graphs brought slightly lower performances, but both grades 9 and 12 showed high levels of performance on interpreting line graphs. Grade 12 students had very low scores on a question on probability. All the results for statistics indicated consistent and orderly growth across the grades.

Consumer Mathematics The inclusion of 18 exercises assessing skills in consumer-related problems was a unique feature of the grade 12 test. The exercises all involved application of mathematics skills rather than knowledge or comprehension. The performances, ranging from 2 percent to 85 percent with a mean of 51 percent, were not high, particularly on items that required knowledge of specialized terminology such as "margin of profit" or "mill rate." Performance levels were highest on comparative shopping, interest calculations, consumer credit, payroll, taxation, and discount

computations, and lower on profit and loss, banking, credit card statements, insurance, investments, and real estate.

The MACOSA mathematics achievement study further analyzed its findings according to sex, school size and, for grade 12, high school mathematics program. The results for these sub-groups were as follows.

Consistent with other research findings, results of the present study showed higher scores for boys on knowledge, comprehension and application items. This sex difference became more pronounced as the grade level increased, resulting in scores that were five or more percentage points higher for boys at the grade 12 level. In Alberta in 1978, boys performed at higher levels in mathematics than did girls.

At each grade level the researcher classified schools as small, medium and large. As with sex, school size showed significantly in accounting for achievement differences, and interestingly, the differences interacted with grade level. For grade 3, the small school (enrolment under 30 students per grade) had the highest performance levels, and quite consistently so over the various content areas. For grade 6 the medium school (enrolment between 31 and 70 students per grade) began to emerge as dominant, and for grade 9 the medium school (41 to 101 students per grade) showed clear dominance. Somewhat surprisingly, school size failed to show any significant difference at the grade 12 level. The large elementary school (over 70 students per grade) quite consistently had the lowest levels of performance. Since most large elementary schools are in large population centers, in general the city elementary schools had slightly lower achievement levels than the rural elementary schools.

At the grade 12 level the test developers designed the mathematics assessment to measure general proficiency, and the test items were not tied to any particular curriculum. However, since one could expect that performances would vary as a result of differences in mathematics programs, the following six groups were used as variables in the analysis of the data: students who had taken one mathematics course, and students who had taken Math 20, Math 23, Math 25, Math 30 or Math 33. Predictably, students who had taken Math 30 had the highest performance levels. Next highest came students who had taken Math 20 or Math 33, and the three remaining alternatives had the lowest performance levels. The greatest differences in performance across mathematical programs occurred on exercises related to number and algebra. Lesser differences occurred on exercises related to geometry, measurement, statistics and consumer mathematics.

Conclusions and Observations MACOSA rated the students' knowledge of number facts and skills in computation (addition, subtraction, multiplication and division) as satisfactory and judged performances in problem solving, geometry, measurement, and consumer mathematics to be generally unsatisfactory.

MACOSA speculated that students tend to do well on topics that are emphasized by the curriculum. For example, generally low scores in geometry might be explained by the recency of the introduction of more geometry to the elementary curriculum and by variations from school to school in the amount of geometry taught at the higher grade levels. Similarly, the recency of the change to the metric system could account for lower performances in measurement. Although the lack of curricular emphasis on consumer education in senior high schools could explain the low scores in this area, MACOSA nevertheless expressed concern because consumer-related skills are important for everyday living in the adult world.

MACOSA members found it more difficult to speculate on reasons for deficiencies in problem solving skills[16], which are required in all areas of the mathematics program. Certainly in the case of word problems, students' reading abilities could be a factor. Perhaps the required reasoning skills involve more steps and more complexities than, for example, the direct application of a single concept to a novel situation. In other words, problem solving may simply be inherently more difficult than many of the other tasks assigned in mathematics. Analysis of the results of this study in terms of students' reasoning abilities might have been enlightening.

MACOSA did not speculate on differences in performance in mathematics according to sex or size of school. Regarding the type of high school mathematics program, MACOSA concluded that Math 30 students should perform best in number work and algebra (as they did) because the Math 10-20-30 program provides the greatest amount of instruction in these areas. Also, since there are fewer differences between programs in the amount of instruction recommended for geometry, measurement, statistics and consumer mathematics, MACOSA did not find it surprising that performance levels for these topics showed lesser differences.

Alberta Science Achievement Study

Purposes The purposes of this study, entitled the *Alberta Science Achievement Study*, were:

1. To investigate current levels of student achievement in science in Alberta at the grades 3, 6, 9 and 12 levels.
2. To provide a data base for future assessments.

Procedures For the grade 3 test the researcher chose 114 items from a variety of sources, and a randomly selected group of 20 primary teachers validated the items. For the grade 6 test the researcher chose 144 items from similar sources, and a panel of 18 upper elementary teachers validated them. A determined effort was made to select items which would reflect the objectives of the elementary science program. There are six content areas in this program — two in physical science, three in biological science and one in earth/space science. Interwoven with these content areas are objectives related to the methods of science, and attitudes toward and knowledge of science and scientists.

[16] Note that problem solving occurs in all content strands of the mathematics curriculum, but applies particularly to the thought level called "application." (See Table 5.)

The tests for grades 9 and 12 were made up of two published standardized tests: the Sequential Tests of Educational Progress (STEP), Series II, and Test of Understanding Science (TOUS), Form Jw. The STEP II measures student achievement levels and the TOUS measures students' knowledge of and opinions about scientists and science as a field of study. For grade 9 the researcher chose the STEP II, Form 3A (with the addition of 15 supplementary items) because it best matched the Alberta curriculum. In grade 12, where the emphasis was to be on more general knowledge of science, both in terms of the content of the high school program and the general aims and purposes of science, the researcher chose the STEP II, Form 2A, because of its relative quality and suitability for gathering the required information. The TOUS, Form Jw, was the only appropriate attitude test among the few such tests which are available.

The Student Evaluation and Data Processing Branch of Alberta Education selected a stratified random sample of schools consisting of 101 schools offering grade 3; 96 schools offering grade 6; 40 schools offering grade 9; and 24 schools offering grade 12.

All tests were administered on May 17, 1978.

Findings Table 6 shows the levels of performance for each grade level and general content area, reported as percentage of correct responses, and the number of items used to test each area.

In grade 3 the performance levels were high on earth science, even though the curriculum does not stress earth/space science; and low on physical science, which includes a number of difficult concepts such as electricity, molecules, and basic energy conservation. Grade 3 students registered a relatively low level of achievement in methods of science investigation despite the fact that science programs have emphasized this area over the past few years. These students demonstrated considerable knowledge in the area of "science as human endeavor", a section of the test designed to check students' perceptions of science and scientists.

The grade 6 curriculum has recently emphasized physical science but, even so, grade 6 scores in physical science were not high. Scores were higher on biological science and life science items and as in grade 3, students at the grade 6 level had high scores on the earth/space science items. Relatively high scores on science methods items probably reflect an increased curricular emphasis on this topic. In grade 6 scores on perceptions of science and scientists (science as a human endeavor) were also relatively high.

Grade 9 students achieved the highest average scores on items dealing with biological and life science, and the next highest average score on physical science. The achievement level in earth/space science was quite low. Scores were high on methods of science while "science as a human endeavor", which included such

topics as aims of science and the skills and aptitudes of scientists, had a low response rate.

Table 6

Science Achievement Levels: Percentage of Correct Responses

Content Area	Grade 3	6	9	12
Physical Science	55.2 (21)*	59.6 (38)	62.6 (17)	55.5 (32)
Biological and Life Science	63.7 (26)	61.8 (45)	69.3 (12)	72.7 (16)
Earth/Space Science	77.2 (6)	70.8 (11)	56.7 (17)	47.9 (3)
Methods of Science	57.4 (41)	64.9 (24)	62.4 (19)	64.8 (36)
Average Student Achievement in Content Areas	59.9 (94)	62.6 (118)	62.2 (65)	62.2 (87)
Science as a Human Endeavor	60.6 (14)	71.1 (14)	49.2 (45)	57.3 (45)
Interest in Science	66.1 (6)	61.2 (12)	—	—

* Number of test items is shown in parentheses.

Grade 12 students achieved at the highest level on biological and life science items and lowest on earth/space science items. Performances were relatively high on methods of science and relatively low on science as a field of human endeavor.

The tests for grades 3 and 6 contained 47 common items, including 18 items drawn from the content areas for grade 3, 18 items drawn from grade 6 content areas, and a further 11 items measuring interest in science and opinions and beliefs about scientists. Table 7 indicates relative student performance on common items.

As expected, grade 6 student performance levels on the common items were substantially higher than grade 3 performance levels. However, the grade 6 performance on grade 6 items was somewhat low.

Relative performance on common items by content area, as shown in Table 8, is not as consistent.

The most substantial differences between grade 3 and grade 6 performances on common items occurred on item clusters dealing with earth/space science, biological and life science, and science as a human endeavor (opinions and beliefs about

scientists), while smaller differences occurred on item clusters dealing with physical science, methods of science and interest in science.

Because both grades 9 and 12 students responded to the TOUS, Form Jw, and to different forms of the STEP II, the researcher was able to make similar comparisons between these grades. Table 9 compares student performances in grades 9 and 12 on the STEP II, for Alberta and the United States. The average response of grade 9 students on the STEP II was 34 out of 50 items, as compared with 32 for their American counterparts. The average score for the grade 12 students was 47 out of 75 items, as compared with 42 for the American norming sample. At both grade levels the higher Alberta performances are statistically significant ($p<0.05$).

Table 7

Performance on Common Items on Science Tests: Percentage of Correct Responses

	Student Performance (%)	
	Grade 3	Grade 6
Grade 3 Target Items j)18)	63.0	77.0
Grade 6 Target Items (18)	49.4	59.1
Interest and Opinion Items (11)	61.4	74.0
Total	57.4	69.4

Table 8

Performance on Common Items on Science Tests by Content Area: Percentage of Correct Responses

	Student Performance (%)	
Content	Grade 3	Grade 6
Physical Science (5)*	62.9	64.3
Biological and Life Science (10)	58.2	74.2
Earth/Space Science (4)	73.1	86.9
Methods of Science (13)	48.4	58.1
Science as a Human Endeavor (10)	57.5	74.9
Interest in Science (5)	61.1	70.0

* Number of test items is shown in parentheses

Table 10 compares performances by students in grades 9 and 12 with the results further broken down to show performance levels for girls and boys.

The TOUS has not been used in Alberta long enough to make comparisons other than those shown in Table 10. As one might expect, grade 12 students scored higher than grade 9 students. In both grades 9 and 12, the girls out-performed the boys. These results suggest that girls understand the nature of science better than do boys.

An examination of grades 9 and 12 results by content area (Table 6) indicated that both groups performed at a high level on the biological and life science items and at a fairly low level on items asking about the nature of science (interest and opinion items). Both grades performed at relatively high levels on items related to scientific methods. Performances on physical science items were relatively high for grade 9 but lower for grade 12.

Table 9

STEP II Test Results: Average Raw Scores

	Form 3A (Grade 9)		Form 2A (Grade 12)	
	Alberta (Spring '78)	U.S.A. (Spring '70)	Alberta (Spring '78)	U.S.A. (Spring '70)
N	2426	2637	2125	2285
Average Score	34 (50)*	32 (50)	47 (75)	42 (75)
Standard Deviation	7	12	12	13

* Total number of test items is shown in parentheses.

Table 10

TOUS, Form Jw, Test Results: Average Raw Scores

	Grade 9			Grade 12		
		Boys	Girls		Boys	Girls
N	2426	1234	1144	2125	1059	1034
Average Score	21.8(45)*	21.1	22.6	25.3(45)*	24.6	25.8
Standard Deviation	6.0	6.1	5.8	6.5	6.7	6.2

* Total number of test items is shown in parentheses.

The researcher also categorized the test items by the thought level required for giving the correct response. Table 11 shows the average performance for three thought levels — knowledge, comprehension and application.

Table 11

Average Performance by Thought Level of Science Items: Percentage of Correct Responses

Grade	Thought Level		
	Knowledge	Comprehension	Application
3	68.3 (16)*	57.8 (24)	67.0 (15)
6	67.4 (38)	57.7 (32)	59.3 (26)
9	64.7 (13)	69.6 (18)	58.6 (25)
12	61.5 (15)	63.1 (15)	58.6 (20)

* Number of items is shown in parentheses.

Elementary students had higher scores than secondary students at the knowledge level, but lower scores than secondary students on comprehension items. Scores on items at the application level were substantially lower than scores at the knowledge level for all grades except grade 3.

Conclusions and Observations MACOSA judged the performance levels in science to be generally satisfactory. The following weak areas were identified: methods of science investigation in grade 3; physical science in grades 3, 6 and 12; and earth/space science and knowledge of science as a human endeavor in grades 9 and 12. MACOSA noted relative strengths across all four grade levels in biological and life sciences, and in grades 6, 9 and 12 for methods of science investigation.

Low scores in grade 3 on physical science items were thought to be acceptable in view of the difficulty of many of the concepts involved. MACOSA speculated that the high scores achieved by elementary students in earth/space science could be attributed to the many children's television programs which emphasize the space theme. At any rate, young children today certainly have a strong interest in the solar system. Similarly, strong performances by elementary students in biological and life science could be attributed in part to television programs produced by organizations such as the National Geographic Society.

Satisfactory student performances on methods of science investigation in grades 9 and 12 perhaps reflect an increased emphasis on laboratory activities in both junior and senior high schools. MACOSA speculated that low scores in grade 12 in the area of science as a human endeavor may have resulted from this topic being considered peripheral to the basic science program.

As noted earlier, most of MACOSA's judgements about adequacy of achievement were made on a subjective basis because no previous norms were established. In the case of science at the secondary level, however, a further measure of adequacy was

available because the science study used standardized tests with United States norms. The average performance for Alberta students in grades 9 and 12 was higher than the U.S. norms established in 1970. There has also been a documented decline in standardized test scores in the United States since 1970; therefore in terms of more recent norms Alberta students are performing at even higher levels than the numbers would indicate (See Table 9, page 43.) On the basis of American norms, then, student achievement in the field of secondary science in Alberta is satisfactory.

Alberta Social Studies Achievement Study

Purpose The purposes of the study, entitled *Report on the Development and Administration of the High School Social Studies Achievement Tests*, were:

1. To establish Alberta norms of performance for high school social studies.
2. To provide a data base for future assessments.

Procedures Initially, MACOSA had hoped to conduct a social studies achievement study similar to the studies done in language arts (reading and writing), mathematics, and science. When a feasibility study (Chandler)[17] was conducted, however, several problems became apparent. Primarily because the Curriculum Branch was at the time conducting an extensive revision of the social studies program, MACOSA decided not to undertake a new social studies achievement project in the knowledge and skills area but rather to report the results of a test which had been developed by the Student Evaluation and Data Processing Services Branch of Alberta Education (SEDPS). This test, which measured performances in social studies at the high school level, had been administered to a sample of Alberta students for norming purposes in June 1977.

The SEDPS test focused on only a part of the range of curricular objectives; that is, only on knowledge and skill objectives. The flexibility of the official social studies curriculum and the procedures involved in developing the objectives and items to be tested leave open two questions: first, whether additional knowledge and skill objectives are present at the classroom level, and second, whether all social studies teachers teach to the objectives identified in the test. Nonetheless, the involvement of over 80 social studies teachers in selecting objectives and developing items, together with the pilot testing of the items, ensure that for the 1971 social studies curriculum the test was as valid as possible.

The test contained 40 items for each course level (Social Studies 10, 20 and 30), or a total of 120 items; Social Studies 30 students answered all 120 items; Social Studies 20 students answered the 80 items for Social Studies 10 and 20; and Social Studies 10 students answered the 40 items directed at that level. The test was constructed to measure two types of performance — mental processes, and knowledge of content

[17] See Chapter 3, p. 15.

areas. The mental processes tested were: recalling; acquiring and organizing; interpreting; and analyzing, evaluating and synthesizing. The content areas are outlined below.

Social Studies 10: participatory government, national unity, poverty, regional disparity, urbanization

Social Studies 20: tradition and change, population and production

Social Studies 30: conflict and cooperation, ideologies, political systems, economic systems

Forty-nine of the above test items dealt with Canadian content. In accordance with the 1971 high school social studies curriculum, most of these items fell in the Social Studies 10 section of the test.

Findings Table 12 provides summary statistics for the total test. Since the length and content of the test differed for each grade level, readers ard advised to be cautious about comparing statistics across grades.

Table 13 shows the average raw scores for grade 12 (Social Studies 30) students on Social Studies 10, 20 and 30 items. Social Studies 30 students scored higher on Social Studies 10 and 20 items than they did on Social Studies 30 items, and they scored higher on the 10- and 20-level items than did students registered in Social Studies 10 and 20.

Figures 2, 3 and 4 show how well Social Studies 10, 20 and 30 students performed in various areas. Figure 2 deals with mental processes, Figure 3 shows performance according to content area, and Figure 4 indicates performance on questions related to Canadian content.

Table 12

Social Studies Achievement Levels

Level of Students	Number of Students	Level of Items	Number of Items	Median Raw Score	Median %	Deviation (raw score)
Social Studies 10	2838	S.S. 10	40	20.1	50.2	7.0
Social Studies 20	2093	S.S. 10 & 20	80	43.7	54.7	12.4
Social Studies 30	1717	S.S. 10, 20 & 30	120	66.1	55.1	17.9

Table 13

Average Raw Scores for Grade 12 Students on Social Studies 10, 20 and 30 Items

Level of Item	No. of Items	Mean
Social Studies 10	40	26.2
Social Studies 20	40	22.4
Social Studies 30	40	19.2
Total Test	120	67.7

Conclusions and Observations MACOSA reviewed only the data related to achievement levels of Social Studies 30 students. Performance levels for these students were judged to be unsatisfactory on the Social Studies 30-level items but satisfactory on Social Studies 10- and 20-level items. Noting that performances by Social Studies 30 students were generally higher than those of 10- and 20-level students, MACOSA concluded that Social Studies 30 students had built upon the knowledge and skills learned at earlier course levels, including in particular the knowledge related to Canadian content, which appears primarily in the Social Studies 10 curriculum.

MACOSA also concluded that deficiencies shown by Social Studies 30 students cannot be generalized across the high school social studies program, but rather must be interpreted in terms of the following: (a) mental process, (b) content area, and (c) available data.

Mental Processes The less satisfactory achievement, when analyzed by mental process rather than by content area (see Figure 2), relates particularly to objectives involving higher mental capabilities, especially the interpretation of information and the processes of analyzing, synthesizing, and evaluating information. Although it might be anticipated that scores in these areas would be lower than in recalling and acquiring and organizing information, scores at the Social Studies 30 level are still below expectations.

Content Areas The content areas (see Figure 3) for which MACOSA judged achievement to be less than satisfactory were participatory government, population and production, conflict and cooperation, political systems, and economic systems. Topics on which MACOSA judged achievement to be satisfactory or slightly above satisfactory were national unity, poverty, and regional disparity.

Available Data

1. The conclusions MACOSA reached about cumulative knowledge and skill rest on the assumption that the results of this cross-sectional study would be duplicated if longitudinal information were available. (Longitudinal information refers to the results of testing the same students over different grade levels in succeeding years instead of different students at different grade levels in the same year.)
2. The impact of the type and quantity of enrolment at the Social Studies 30 level presents an unknown factor. Presumably the average level of achievement for Social Studies 30 students might be greater than that for Social Studies 10 and 20 students because the 30-level group tends to include the more academically successful students. Credit in Social Studies 30 is not required for the high school diploma; and of course students who drop out of school or fail prerequisite courses are eliminated from the Social Studies 30 group.
3. MACOSA speculated that the lower achievement of Social Studies 30 students on the 30-level items could be explained by the apparent curricular leap to a considerably higher order of complexity in concepts and skills at the Social Studies 30 level.
4. A lack of information about performance levels in social studies for elementary and junior high school students detracts from MACOSA's confidence in the conclusions it has reached about progressive achievement. Such information might have influenced conclusions about the social studies in general or about progression in the high school program, but the possible extent and direction of such influence is not known at this time.

Several peripheral questions arose during the discussion of achievement levels in social studies. MACOSA noted that the difficulty test developers experienced in preparing a test for province-wide use was largely related to a lack of specified common elements in the curriculum. If such tests are to be given in the future, a clearly specified core curriculum would be desirable. Also, because of recently approved changes in the Alberta curriculum for social studies, the curricular validity of the present SEDPS high school achievement test for the social studies may be in doubt. Finally, MACOSA hypothesized that the performance levels on the present test may bear a substantial relationship to how successfully the 1971 social studies curriculum was implemented in the schools.

Figure 2 Achievement Levels of Social Studies 10, 20 and 30 Students by Mental Process: Percentage of Correct Responses

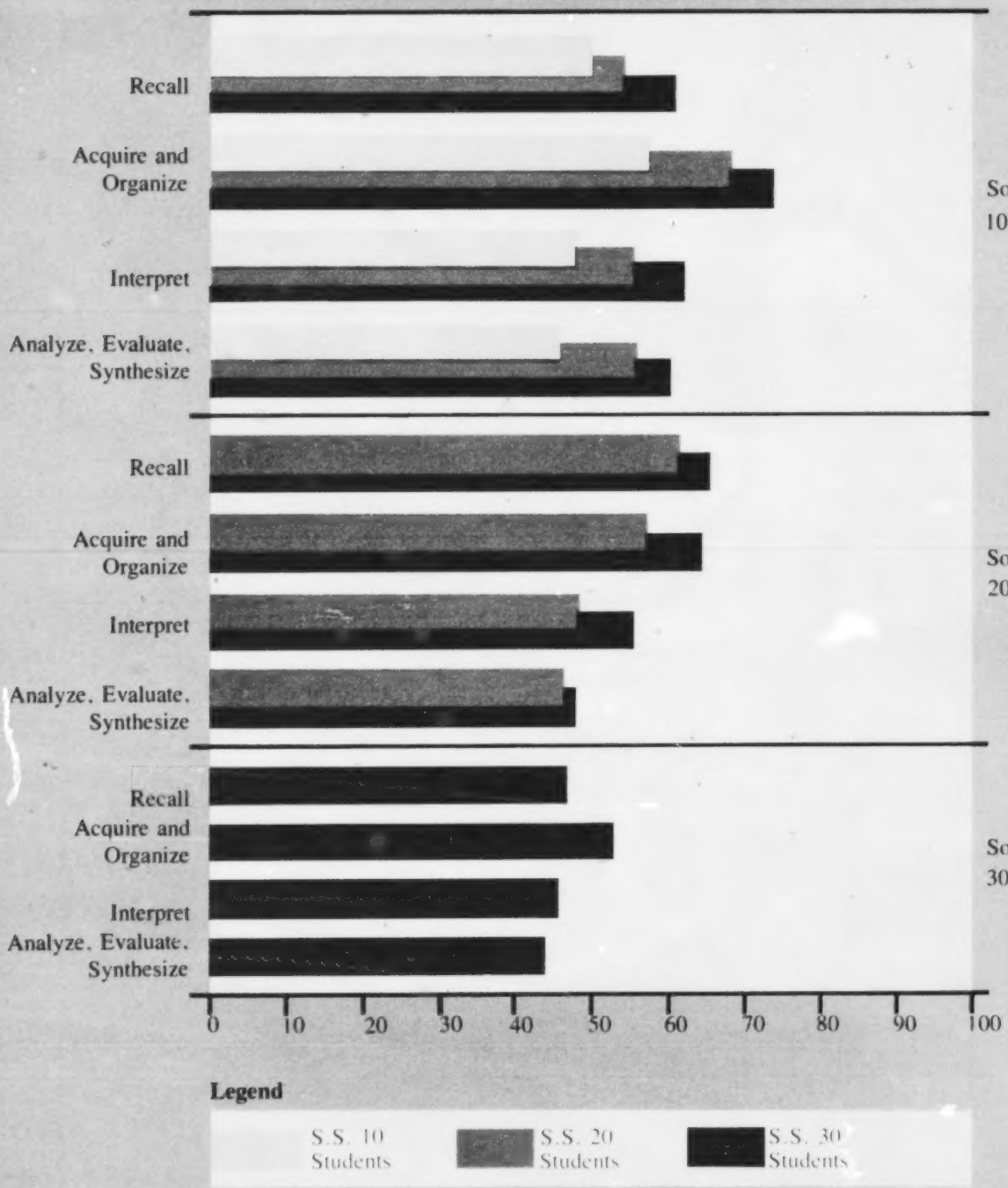


Figure 3 Achievement Levels of Social Studies 10, 20 and 30 Students by Content Area: Percentage of Correct Responses

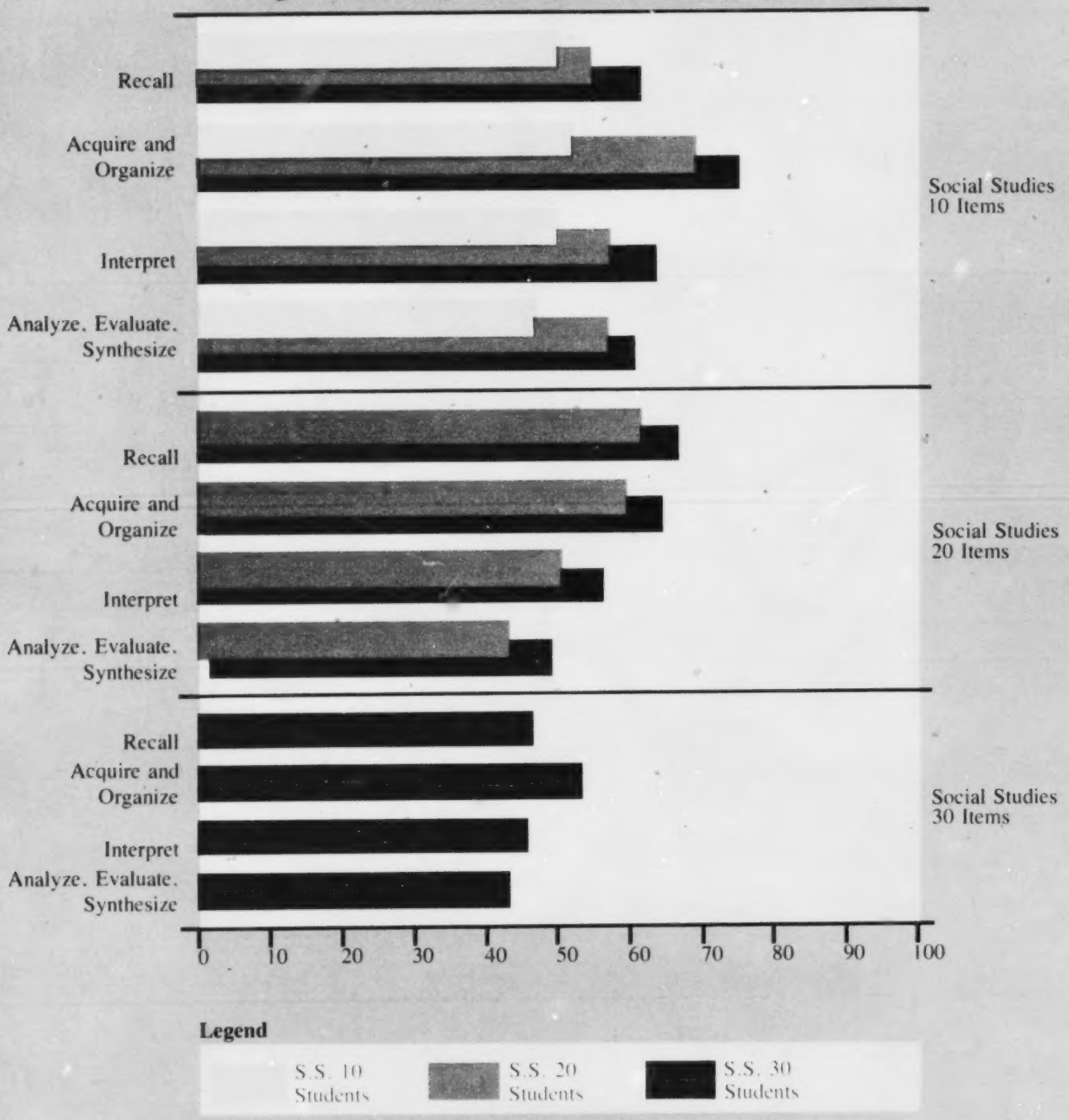
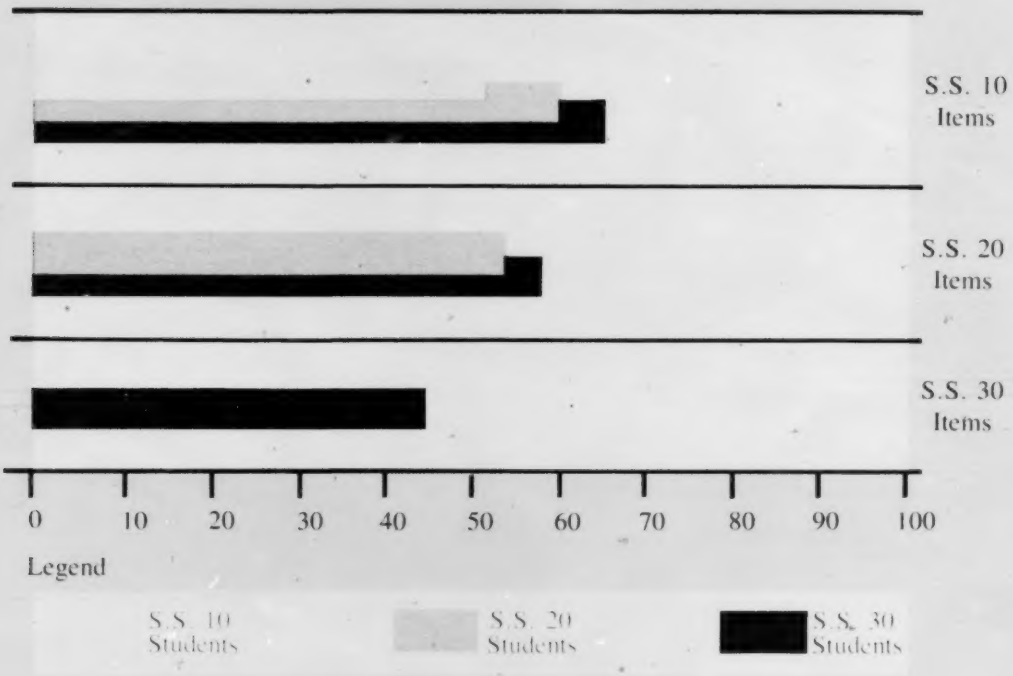


Figure 4

Achievement Levels of Social Studies 10, 20 and 30 Students on Canadian Content Items: Percentage of Correct Responses



CHAPTER 6

Test Development Studies

On reviewing the Alberta curriculum, MACOSA noted that a fair amount of emphasis is placed on the development of oral communication skills and on students' attitudes and values, but that little research has been done regarding the development of tests to measure achievement in these areas. MACOSA commissioned four studies to explore the possibility of developing and validating such tests for Alberta: a two-part study of listening and speaking; and three attitude studies, including attitudes towards the world of work, attitudes towards school subjects, and an assessment of affective objectives in the social studies.

Listening and Speaking Study

Purpose The purpose of the listening and speaking study[18] was to design and/or adapt tests for assessing the achievement of selected objectives in listening and speaking in grades 3, 6, 9 and 12.

Procedures The MACOSA listening and speaking study was conducted in two stages. Because the Alberta Education curriculum guides stated objectives for listening and speaking programs in rather general terms, a researcher first developed a list of measurable objectives. Then, after the MACOSA listening and speaking steering committee and the language arts coordinating committee of the Curriculum Branch of Alberta Education had established priorities among these objectives, a second group of researchers developed and validated tests of listening and speaking appropriate for students in grades 3, 6, 9 and 12.

The second group of researchers reviewed available tests of listening and speaking but, finding no usable tests or items, they constructed all new items. The tests were piloted and field tested, and a final test was given to a sample of Alberta students. All testing was done under controlled but natural conditions, using a tape-recorded format. Both tests were kept to a length of one hour or less.

A review of the literature and the results of the pilot and field testing helped the researchers to identify a number of general principles, both practical and theoretical, about the testing of oral communication. They decided that the specific objectives upon which the tests were based should include knowledge, skills and attitudes; take into account the various audiences, situations and purposes related to speaking and listening activities; and recognize the various levels of language usage (casual, informal and formal). The tests were designed to evaluate as fully as possible the range of objectives described above. Both the listening and speaking tests attempted to evaluate the expressions and reception of three kinds of oral language — spontaneous spoken language, prepared spoken language, and written language read aloud. The tests also evaluated a variety of oral language factors, including vocal, verbal and non-verbal factors (voice production, vocabulary, syntax, fluency, variety, organization, content).

[18] The listening and speaking study produced two reports, entitled *Listening and Speaking Objectives* and *Test Development for Assessing Achievement in Listening and Speaking*.

Conclusions and Observations The listening and speaking test project was a pioneering venture in a new area of evaluation; consequently, the tests are still in the developmental stages. In view of the satisfactory results of this study, however, the researchers are optimistic about continued development and revision of procedures to assess achievement levels in these two important elements of the language arts program.

The tests are designed to assess students' mastery of certain skills as described in the program objectives rather than to compare individual students with each other. In general, the tests are adequate for measuring the achievement of groups of students with respect to particular objectives and for evaluating the success of an instructional program.

The test results would be more useful if provincial norms and performance standards could be developed and made available. Further analysis and revision would also provide more information about the tests and increase the users' confidence in them.

Attitude Studies

MACOSA commissioned two studies to develop and validate tests for measuring student attitudes towards the world of work and school subjects, and a third study to develop test models for assessing affective objectives in the social studies. MACOSA was motivated to undertake the development of attitude tests because: 1) a number of the goals of basic education for Alberta are of this type; 2) these goals are important and merit attention; 3) commercially available tests do not measure in this area, or lack technical adequacy, or are judged to be inappropriate for Alberta schools; and 4) studies are needed to determine whether testing in this area (outcomes of schooling related to attitudes and values) is feasible.

The time lines for the studies were compressed; work in this area was not begun until May 1978. The researchers for the studies of attitudes towards the world of work and school subjects used conventional procedures in developing the tests and determining their reliability and validity. The use of factor analysis[19] and test-retest reliability procedures, for example, provided valuable insights about the nature of the tests and defensible uses for them. Both groups of researchers reviewed the literature in their areas of focus and then proceeded beyond the work of previous researchers.

Norms for these instruments which would provide benchmarks for monitoring trends and interpreting future results have not yet been established.

Study of Attitudes Towards The World of Work

Purpose The purpose of this study, entitled *Development of Scales on Attitudes Towards the World of Work*, was to develop and validate a test to assess the attitudes of Alberta students towards the world of work in the following areas: attitudes

[19] Factor analysis is a procedure used to identify factors that account for the correlations among items or test scores. See "*School Subjects Attitude Scales*," p. 56.

towards earning a living, perceptions about employer expectations, perceptions about available opportunities, relevance of school preparation for employment, and characteristics of desirable jobs.

Procedures The researchers developed two pilot tests on the basis of a literature review and with reference to the five areas of concern described above. Each test item consisted of a statement about the world of work, along with a five-point rating scale which ranged from "Strongly Disagree" to "Strongly Agree". These tests were administered to samples of students in grades 8 and 11.

Two revised tests, developed on the basis of the piloting, provided for assessment of the following 15 opinion scales: preparation by the school, interest and variability in jobs, diligence, laziness, job security, positive employer characteristics, independence, money, ambition, locus of control, confidence in succeeding, negative employer characteristics, social relations, attitudes towards unemployment, and general attitudes towards earning a living.

Students in grades 9 and 12 from eight separate geographic locations in Alberta wrote the revised tests.

To check for validity, the researchers compared groups of students according to their level in school (junior or senior high), sex, school program, and job aspirations; and on the basis of whether they had held full-time or part-time jobs, and whether they were doing well in school. The high school groups were also compared with 77 students enrolled in an academic upgrading program at NAIT and 48 students enrolled in a similar program at the Alberta Vocational Centre in Edmonton. Test-retest reliabilities were determined for the scales using a five-week interval between test administrations.

Most tests of this kind are designed to be independent of a particular social and economic context. These MACOSA tests, however, seek to inquire about students' attitudes towards work in the very context they are likely to enter — the Alberta labor force. Another difference from most previous studies is that these tests use opinion scales which aggregate the views of groups of students without attempting to ascertain why these views are held or provide psychological analyses of individual students.

Findings Analyses of student opinions indicated that the 15 scales were relatively independent of each other and structurally valid.

The sample (which was not necessarily representative of Alberta students on the whole) gave the overall impression that the work ethic is deeply entrenched in the expressed opinions of Alberta students. Typical views included: "There are jobs available for those who want them"; "A person should feel a little ashamed for doing a sloppy job"; and "One of the most important things about a job is to know that you are doing the best you can". The sampled students appeared confident in

their ability to succeed as evidenced in their responses to items like, "I will be able to handle the next step in my training"; and "When I finish my education I'll be able to get the kind of job that I want".

Approximately one-third of the students responded to a section of the test which requested written comments. The most common opinion expressed was that the level of vocational counselling in the schools is inadequate.

In nearly every case the test results appeared to be consistent with what might be expected from the designated groups. For example, students who were doing well in school saw interest and variability as being important in a job, they seemed to value diligence, they saw employers in a positive light, they sought independence, money was a lesser concern, they saw themselves as being able to control their job situation, they were far more confident than students who were not as successful in school, and they seemed to value the social aspect of a job. Girls gave lower ratings than did boys to job security, independence, ambition and money. Girls rated interest and variability in a job higher than did boys, were more negative about laziness, and gave lower ratings to "negative employer characteristics". The social aspects of work seemed more important to girls.

Satisfactory test-retest reliability results were obtained. (The reliability coefficients ranged from 0.46 to 0.70 with a median value of 0.61. Only one of the 15 scales showed a reliability coefficient below 0.50.)

The data indicate, therefore, that the scales on Attitudes Towards the World of Work are sufficiently valid and reliable for use in securing information about groups of students in grades 9 through 12. The tests also have face validity; that is, the tests appear to measure what they are intended to measure. Educators who administered these scales showed considerable interest in the results, particularly in the students' support of the work ethic, but also in the many other aspects of employment which these scales tapped.

Conclusions and Observations The researchers concluded that these scales have a number of potential uses, such as collecting student opinions over an extended period of time or among various subgroups, and as a stimulus for classroom discussions of vocational and guidance matters. The results of this test could also make teachers aware of curricular needs related to the world of work.

School Subjects Attitude Scales

Purpose The purpose of this study, entitled *Development of School Subjects Attitude Scales*, was to develop and validate for use in Alberta schools a test of students' attitudes towards school subjects.

Procedures The school subjects attitude scales used a semantic differential approach. A semantic differential test asks students to indicate the meaning they attach to a given concept (in this case a school subject) by recording it on a

five-point rating scale between carefully selected bi-polar adjectives such as "good-bad", "hard-soft", or "fast-slow". Figure 5 illustrates the format of each item on the scale.

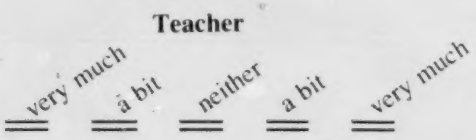


Figure 5. Sample Item from Semantic Differential Test

The researcher assembled a list of more than 100 bi-polar adjectives from various sources, including other attitude scales, and selected from this list 24 pairs for a long form of the test and 12 pairs for a short form. The selections were made after discussions with teachers and students and analysis of trial forms. The scales, which were directed towards students in grades 5 and 8, were designed to measure attitudes of groups of students.

As a validity check, the researcher hypothesized the direction of differences between groups with known characteristics and tested whether the scales discriminated between these groups. Sex differences, cultural differences, and differences related to type of school and aptitude or intelligence were analyzed. Test-retest reliability was also checked.

Findings Field trials of the attitudes towards school subjects scales indicated that this test is reliable and valid; however, pilot testing has been necessarily limited. The scales' remarkable face validity suggests that they would have credibility with school personnel.

Factor analysis revealed three major dimensions in the scales: "evaluation" (general feeling towards the subject), "usefulness" of the subject, and "difficulty". Comparisons of the different groups produced the expected results, indicating that the scales are valid. For example, the scales indicated that males had a more positive attitude towards science than did females. (It is interesting to note that correlations between attitudes and achievement were generally low or not significant and that correlations between attitude and intelligence were not significant.)

Conclusions and Observations The study concluded that Alberta norms should be established for these scales. One use for the scales after norming might be program evaluation, since the scales could assess student attitudes before and after a program change.

Assessment of Affective Objectives in the Social Studies

Purposes The purposes of the study, entitled *Development of Models for Assessing Affective Objectives in the Social Studies*, were:

1. To identify and further specify affective objectives in the social studies.
2. To develop models of tests to assess achievement of these objectives.

Procedures The study was directed towards the social studies programs for grades 6 and 12. The researcher first developed a chart which classified the many affective objectives in the social studies according to categories such as feelings, beliefs, attitudes and values. Since the range of objectives was very broad, the researcher focused on citizenship education, which has traditionally received a major emphasis and is still a high priority in the current social studies curriculum. Through the cooperation of the Curriculum Branch of Alberta Education and its coordinating committees the objectives chosen for use in test development were carefully referenced to Alberta Education curriculum guides and to the *Goals of Basic Education* statement.

The next step involved various committees of teachers who constructed models and appropriate test items for assessing progress towards the selected objectives. The study produced 8 test models for grade 6, and 22 test models for grade 12. The test models can be administered separately or in sets. Since the scope of the study did not allow for field testing, the validity of these models has not yet been determined.

Conclusions and Observations MACOSA concluded that validated versions of these test models could assist teachers in evaluating achievement of the affective objectives for citizenship education in social studies programs. They could also be used to monitor attitudinal changes among students over time and to determine the needs and interests of students.

MACOSA also noted that clear and detailed specification of curricular objectives in the affective area would greatly assist in the development of these tests. (See Recommendation #2, page 66.)

CHAPTER 7

Grade 12 Departmental Examinations

In response to growing public concern, the Legislative Assembly of Alberta passed a motion on October 19, 1976 requesting that a study be conducted to examine the effect of the withdrawal of grade 12 departmental examinations on student achievement. The task was later assigned to MACOSA, which was established at about the same time (October 1976). The resulting *Alberta Grade 12 Examination Study* was the first study commissioned by MACOSA.

Alberta Grade 12 Examination Study

Purpose The purpose of the study was to determine whether the change to non-compulsory grade 12 departmental examinations had affected the quality of education in Alberta.

Procedures The study approached the problem on the following three fronts:

1. What are the current grade 12 student evaluation policies and practices at the school system and individual school levels across the province?
2. What changes have taken place over the last five years in the distribution of marks awarded by the schools in grade 12 subjects?
3. What do the stakeholder publics think has happened to the quality of education since accreditation; what changes do they think can be attributed to the change in grade 12 examinations; and what do they think ought to happen now?

The researcher surveyed superintendents in order to answer the first question, and the Student Evaluation and Data Processing Services Branch of Alberta Education compiled the necessary data from their records to answer the second question. Regarding the third question, the researcher used a direct mail questionnaire to gather opinions. A total of 6800 individuals from ten stakeholder groups were selected to respond. The average rate of return of the questionnaires was 65.4 per cent, ranging from a low of 34 per cent return from public sector employers to a high of 100 per cent from the 52 schools selected at random from a total of 290 Alberta schools offering instruction at the grade 12 level in 1977. The public-at-large sample consisted of 3000 names chosen using a combination random and fixed factor (geographic area) sampling design.

In order to minimize limitations imposed by this method of opinion gathering, personal interviews were conducted with a 2 to 3 per cent sample of respondents from each stakeholder group. The interview was structured to perform a validation function as well as to determine whether the questions were clear to the respondents, whether respondents had additional ideas regarding some of the main topics, and what reasons respondents had for their answers to certain key questions.

Findings Regarding current evaluation policies, the study found that few systems have no local policy on student evaluation — only about 10 per cent of students were in systems without a policy. Most systems had adopted modified forms of a

policy originated by the Lethbridge school systems. This policy hinged on means for ensuring fair and just evaluation (as also specified in Alberta Education's accreditation guidelines). However, a wide gap existed between policy at the school board level and practice in the schools. The apparent lack of consistency in application of policy in actual school operations, within the same system specifically and throughout the province generally, was perceived by superintendents as a major obstacle in maintaining credibility with the public. Superintendents also identified a number of other concerns and problem areas which, taken as a whole, would indicate that there existed an urgent need to examine ways of improving policy implementation procedures.

A review of high school records showed that since compulsory examinations were dropped in 1973 there had been a general increase in marks awarded by school staffs in all academic subjects except Chemistry 30X and Physics 30X. The typical trend of marks since the 1970's was a sharp increase in all regions of the province to 1973, followed by a more gradual increase up to 1975. A slight drop was evident in 1976 and 1977. A continuation of this trend to 1980 would likely result in a return to the five-year average established under the former system of external examinations.

The results of the questionnaire survey indicated that most people thought that changes had occurred in the quality of education in the past five years. People from outside the professional educational community tended to feel that the change had been for the worse. These people perceived a lack of common standards across the province and felt this lack was a problem. They viewed the examinations as comparable guideposts, and nothing appeared to have replaced them in this function. Educators, on the other hand, were less enthusiastic about a return to a compulsory departmental examination system. Some teachers were pessimistic about a simple solution to this complex problem.

More specifically, the survey indicated that:

1. Respondents thought that mathematics and science achievement may have improved or remained constant but that achievement in English had decreased.
2. There was general but not universal feeling that some change was needed, but the groups responding were split as to whether compulsory departmentals should be re-established. Adults generally favored reintroduction, but subcategories of adults varied in opinion.
3. There was overwhelming agreement that final marks should be based on the entire year's work and not on a single examination (but a final examination could be a significant part of the final evaluation).
4. Respondents did not think that multiple choice testing alone could adequately test all the different kinds of student achievement.

5. Respondents were favorable to evaluation procedures measuring both knowledge and attitudes, but interviews indicated uncertainties as to what attitudes should be assessed, and how.

The researcher noted that a further problem with using the pre-1973 form of departmental examinations was a tendency to treat non-examination subjects as second class subjects. Arguments that seemed favorable to the return of examinations were the maintenance of comparable standards across the province and certification of a known level of achievement. Interestingly enough, even though the majority of respondents agreed with the latter reason, there was nothing in the pre-1973 grade 12 departmental examinations that guaranteed this. Although students could be compared within a year, there was no guarantee of comparability across years. (Examinations of differing degrees of difficulty, and sometimes having different curricular emphases, were administered each year and the results were scaled each year so that the same proportion of students received As, Bs and so on with some variation in the different subject areas.)

The arguments surrounding the departmental examinations are not simple ones. Some of the intentions that surrounded their withdrawal may not have been realized; others have. Societal change has contounded the issue, making it difficult to reach a clear perception of what has happened.

Conclusions and Observations In its report of November 1977, the study concluded that the answer to the question from the Legislative Assembly depends on the interpretation of the question, including what constitutes quality.

1. If "quality of education" means the amount students actually know, there is no data available for answering the question, given the difficulties of measurement and given the changes in the curriculum.
2. If quality of education means standards of achievement and these standards are equated with the proportions of A-F marks awarded, and assuming the inherent abilities of Alberta students are much the same from year to year, then in more jurisdictions and subjects more 'A' grades are awarded now than previously, and this "inflation" can be inferred to be a result of dropping mandatory examinations.
3. If quality of education is interpreted as people's perceptions about quality, then change has occurred, but groups are split as to the direction of change. Persons most closely associated with education tend to view the change as more positive than does the public at large. Changes in society, rather than the withdrawal of departmental examinations, were more frequently identified as a cause of change.

4. Given other changes in society, a return to departmental examinations will not return education to pre-1973 conditions. However, something needs to be done because of the perceived decline in achievement as opposed to the rise in marks, the general decline in public confidence in education, and the continuing existence of room for improvement in the quality of education.

Conclusions of MACOSA In October 1977, when the *Alberta Grade 12 Examination Study* was released, MACOSA recognized that the Alberta public was seriously concerned about educational quality but that at the same time some groups were reluctant to return to the pre-1973 form of departmental examinations. At that time MACOSA recommended on an interim basis:

> That the Minister of Education extend the existing accreditation policy for a minimum of three years, subject to review after that period, and further that the necessary steps be taken to enhance the effectiveness of the existing policy with respect to student evaluation.

Subsequent Events[20]

1. After MACOSA submitted the above recommendation to the Minister of Education, the Minister announced that Alberta Education would intensify its supervision of school evaluation policies and procedures to reduce major discrepancies among systems and schools, and that the Curriculum Branch of Alberta Education would continue, as the Curriculum Policies Board had recommended, to more precisely specify learning objectives in programs of study and curriculum guides.
2. In 1978 MACOSA received as information a study conducted at the University of Alberta (Mehra, 1977) which found that teacher-assigned marks in university entrance subjects are slightly better predictors of success in university programs than were the departmental examination marks assigned before 1973.

[20] Subsequent events regarding the grade 12 departmental examination question are discussed in greater detail in Chapter 8, under Recommendation #11.

CHAPTER 8

General Conclusions

General Conclusions and Recommendations

MACOSA commissioned five preliminary studies, five achievement studies and eight other studies related to student achievement. The researchers submitted data, conclusions and recommendations to their steering committees, and both results interpretation panels and steering committees made judgements concerning standards of performance for the achievement studies. On the basis of information from all of these sources, MACOSA arrived at some general conclusions about student achievement in Alberta and suitable evaluation procedures. These conclusions, which appear in greater detail throughout the report, are summarized here.

1. **Alberta Grade 3 Achievement** Performance levels of a sample of Alberta grade 3 students on standardized tests administered in 1977 were assessed. Comparisons were made with data from a similar study conducted in Edmonton public schools in 1956 and between rural and urban schools in 1977. From the findings of this study, MACOSA concluded that skill levels in reading, language and arithmetic have been maintained for Edmonton students over the 21-year period. The 1977 skill levels on all tests were about the same for rural and urban groups, an apparent improvement in the situation in that many people in the past believed that student achievement was lower in rural schools, a belief that had some support from the research data available.

2. **Reading (Grades 3, 6, 9 and 12)** Performances in reading were generally satisfactory at all four grade levels tested, and the performance in phonics at the grade 3 level was particularly satisfactory. Alberta students appeared to be improving or at least maintaining their reading skills as they progressed through the grades. Students showed weaknesses, however, in using prefixes and suffixes to obtain word meanings in grade 6; in identifying directly stated details on maps and charts in grade 9; and in understanding implications in selected reading materials in each of grades 6, 9 and 12.

3. **Writing (Grades 3, 6, 9 and 12)** Performances were weak in all four grades on the short answer section of the writing test, which assessed abilities in capitalization, spelling, punctuation and sentence structure. Performances were stronger on the longer writing assignment, which asked students to write a story in grade 3, a story or description in grade 6, an exposition in grade 9, and an argument in grade 12. On the writing assignment all students achieved high scores for keeping on the topic and writing legibly. Alberta students appeared to be more successful at using their own language in writing than at finding errors in or manipulating a structured assignment. Although performance on the writing assignment was generally satisfactory, there was some concern about the fact that 26 per cent of grade 12 students performed below the expected functional level. (Students whose writing was rated as "func-

tional'' were able to communicate at a basic level but needed instruction in some aspects of writing skills.) MACOSA judged performances in writing to be satisfactory in grades 3 and 9, marginally satisfactory in grade 6, and unsatisfactory in grade 12.

4. **Mathematics (Grade 3, 6, 9 and 12)** Performance in number facts and computation (adding, subtracting, multiplying and dividing) was satisfactory at all grade levels, but performances in problem solving, geometry, measurement and consumer mathematics were unsatisfactory. MACOSA speculated that the difficulty with problem solving might be attributed to weaknesses in students' reading abilities as well as to the complex reasoning skills required. (No firm evidence is available to support this conclusion.) Weaknesses in geometry, measurement and consumer mathematics might be related to the lesser emphases these areas receive in the Alberta mathematics curriculum. Geometry studies at the elementary level have only recently been increased. The metric system of measurement is relatively new to Alberta school programs, and consumer mathematics, although considered an important skill for daily living, receives little or no attention in the senior high school.

   This study found that boys performed better than girls on all aspects of mathematics and this difference became more pronounced as the grade level increased. School size also showed a relationship to achievement differences, and interestingly, the differences varied according to grade level. Grade 3 students in small schools (under 30 students per grade) performed best; in grade 6 the medium-sized school (31 - 70 students per grade) began to emerge as dominant; for grade 9 the medium school (41 - 101 students per grade) showed clear dominance; and at the grade 12 level school size failed to show any significant differences in achievement. The large elementary school (over 70 students per grade) quite consistently had the lowest levels of performance. Regarding the type of high school mathematics program, students who had taken Math 30 had the highest scores. It is not surprising that the Math 30 students performed best in the areas where they had received the greatest amount of instruction (number work and algebra) and that they were fewer differences among students in the different mathematics programs in areas where the amount of instruction is more equivalent (geometry, measurement, statistics and consumer mathematics).

5. **Science (Grades 3, 6, 9 and 12)** Performances in science were generally satisfactory at all four levels. Weaknesses were evident, however, in the following areas: methods of scientific investigation in grade 3; physical science in grades 3, 6 and 12; and earth/space science, and knowledge about science and scientists in grades 9 and 12. All four grades showed relative

strengths in biological and life sciences, and grades 6 9 and 12 showed strengths in methods of scientific investigation. The low scores in physical science at the grade 3 level could be attributed to the difficulty of the concepts involved. Satisfactory performances on scientific processes in grades 9 and 12 probably reflect an increased emphasis on laboratory activities in junior and senior high schools.

In the case of science at the secondary level, an additional measure of adequacy was available in that standardized tests with United States norms were used. The average performances for Alberta students in grades 9 and 12 were higher than the U.S. norms established in 1970.

The science study also showed that in grades 9 and 12 girls outperformed boys on questions measuring knowledge of and attitudes towards scientists and science as a field of study.

6. **Social Studies (Grades 10, 11 and 12)** Because the Alberta social studies program was under revision when MACOSA conducted its achievement studies, only performances at the high school level were assessed.

   Social Studies 30 students scored higher on Social Studies 10 and 20 items than on Social Studies 30 items. Also, Social Studies 30 students scored higher on the 10- and 20-level items than did students currently registered in Social Studies 10 and 20. Performances of Social Studies 30 students were considered unsatisfactory on the 30-level items but satisfactory on the other items. Students did less well on questions which required skills in interpreting information or in analyzing, synthesizing, and evaluating information than on questions requiring the use of simpler processes like recalling or acquiring and organizing information. Achievement was unsatisfactory in the content areas of participatory government, population and production, conflict and cooperation, political systems and economic systems; and satisfactory in the areas of national unity, poverty, and regional disparity.

   MACOSA observed that (1) if province-wide achievement testing in social studies is to be successful, there must be a clearly specified core curriculum; and (2) performance levels on the MACOSA social studies test may bear a substantial relationship to how successfully the 1971 social studies program was implemented in the schools.

7. **Grade 12 Examinations** The findings of the *Alberta Grade 12 Examination Study* led MACOSA to conclude that Alberta Education should not reinstitute the departmental examinations in their pre-1973 form. However, in view of the perceived decline in achievement as opposed to the rise in marks, the general

decline in public confidence in education, and the continuing need for improvement in the quality of education, MACOSA recognized that appropriate information about student achievement should be made available and that action should be taken to facilitate the achievement of fairness and equity in the awarding of marks.

MACOSA's recommendations in this regard appear later in this chapter. (In 1977 when the *Alberta Grade 12 Examination Study* was released, MACOSA made an interim recommendation that the existing accreditation policy be extended for a minimum of three years, subject to review after that period, and that the necessary steps be taken to enhance the effectiveness of the existing policy on student evaluation.)

8. **Factors Affecting Student Achievement** MACOSA's testing program could not take into account factors affecting student achievement because these factors are difficult to measure and interact with each other in complex ways. MACOSA also decided that such information was not needed for its assessments, which dealt with data at the provincial level only and did not evaluate individual students, schools or school systems.

   However, recognizing that home environment, aptitude and other non-school factors can be as significant as instructional factors (or even more so) in the evaluation of student achievement, MACOSA commissioned two reviews of the literature to further analyze the research in this area. These reviews found that the research tends to be inconclusive in terms of any consistent effect of environmental or school factors on student achievement.

9. **Listening and Speaking** Members of MACOSA are hopeful that the tests developed for listening and speaking for grades 3, 6, 9 and 12 can be made reliable and valid through continued development and revision and that Alberta norms can be established.

10. **Attitude Studies** The researchers judged that the tests they developed to assess attitudes towards the world of work and school subjects were reliable and valid. MACOSA observed that these tests could be quite useful to schools after further refinement and the establishment of norms, and agreed that a project which developed models of tests of the citizenship education objectives in the social studies program should be continued.

Recommendations

Having received all the information which was requested, including test results, reviews of research, judgements of panels and steering committees, overviews of possible evaluation systems and structures, and data from various other studies, the Ministers' Advisory Committee on Student Achievement makes the following recommendations, which have been grouped into three areas.

I. A Proposed Assessment Program

The Minister's Advisory Committee on Student Achievement recommends:

1. THAT MACOSA-type tests be administered periodically at selected grade levels, but with sampling plans extended to permit generalization of results to school system populations.

This testing program would serve the following purposes: feedback to publics, including reports on how local systems meet particular goals and objectives set by the province; and decision making regarding maintenance and improvement of the quality of instruction, including local system needs vis à vis the particular provincial objectives tested.

Details of a proposed testing program of the type recommended are given in Table 14. This suggested program involves cyclical testing at three-year intervals for each subject assessed, with no more than three subjects to be tested in any given year. Testing would occur mainly at the grades 3, 6, 9 and 12 levels but variation would be possible. In such a testing program a relatively few selected high priority objectives would be used in the development of the test for each subject area.

This testing program could serve school systems by using provincial resources which are to a large extent already available. Otherwise, school systems would incur considerable costs for developing and administering their own tests. The program would provide information about the distributions of scores at both provincial and school system levels to all interested groups and to school boards, administrators and teachers. Provincial and local averages and the variability of distributions would also be communicated.

2. THAT the proposed assessment program test curricular objectives selected from all domains (knowledge, attitudes and skills) and thought levels, and not only from the most easily tested areas.

Assessment programs of this type can unduly influence the instructional emphases of school programs because educators tend to concentrate on those objectives which are tested and place less stress on areas that are not tested. If the tests are designed to measure achievement of a wide variety of curricular objectives, this problem is less likely to occur.

The MACOSA tests attempted to measure achievement of a wide range of high priority objectives. For example, the writing test evaluated students' abilities in writing original compositions as well as aptitude for identifying errors in punctuation; and the science test assessed students' attitudes towards science as well as knowledge of certain concepts.

MACOSA noted the difficulties experienced in testing attitudes and values because of the absence of a clear and detailed specification of curricular objectives for affective learning.

3. THAT the tests developed by MACOSA investigators be revised as required, and that those tests which have not been given be normed in order that they may be used in subsequent assessments of student achievement.

Tests which have not yet been normed include the Listening and Speaking tests, the Scales on Attitudes Towards the World of Work, and the School Subjects Attitude Scales. Of these tests, only the Scales on Attitudes Towards the World of Work are presently ready to be administered. The other tests require varying amounts of revision. Norming of these revised tests would take place only if the tests are found to be valid and reliable.

Table 14

Specific Features of a Proposed Provincial Testing Program

Month/Year	Grades	Curriculum Area	Comments
May, 1980	3,6,9,12	1. Science	1. Use 1978 tests, possibly with some additions.
	3,6,9,12	2. Listening and Speaking	2. Provincial norming administration.
May, 1981	3,6,9,12	3. Mathematics	3. Use 1978 tests, possibly with some additions.
	6,9	4. Attitudes Towards School Subjects	4. Spring norming administration for grades 6, 9. Administer to grades 11 or 12 if validated for use in senior high school.
	9,12	5. Attitudes Towards World of Work	5. Alternatively, may be used with grades 8 and 11.
May, 1982	3,6,9,12	6. Reading	6. Use 1978 test, possibly with some additions.
	3,6,9,12	7. Writing	7. Use 1978 tests.
	6,12	8. Social Studies (Civics Education)	8. Permits time for further test development (e.g., other aspects of affective domain) and for additional grades (e.g., grades 3, 9).
May, 1983		Repeat 1980 portion of the cycle	Use expanded test in listening and speaking — 1978 version with additions.

4. THAT at some appropriate time the MACOSA tests in Reading and Writing, and Listening and Speaking be used to investigate relationships among various aspects of the language arts in order to determine recommendations about instruction.

At the present time there is no firm evidence regarding the above-mentioned relationships, particularly those that might pertain to an integrated language arts curriculum. The MACOSA listening and speaking study, which was limited in scope, found that good speakers are not necessarily good listeners, and vice versa. Further research in this area could result in modification of instructional methods or even of curricular objectives. For example, further research might show that teachers should conduct separate activities directed towards teaching each of listening and speaking.

5. THAT the work initiated by MACOSA in assessing the affective objectives of the social studies be continued.

MACOSA commissioned the development of test models for assessing attitudes and values in the area of citizenship education. This recommendation endorses piloting and field testing of these models.

6. THAT the proposed assessment program and the present high school achievement tests be integrated.

Although in a number of cases the MACOSA tests for grade 12 and the high school achievement tests are quite different, MACOSA observed that duplication of effort should be avoided where the reasons for testing are similar and the tests themselves have similar content. This procedure could also reduce costs.

7. THAT validation procedures for tests which include items outside the specified core curriculum include the involvement of a cross-sectional sample of teachers and community representatives.

This recommendation would apply, for example, to the grade 12 mathematics test which contains items on consumer education even though that topic does not appear in the grade 12 mathematics curriculum. In such cases information about what students should be expected to know is not available from the program of studies and must be obtained from knowledgeable educators, members of receiving groups such as post-secondary educational institutions, business and labor, and members of the general public.

8. THAT where appropriate, results interpretation panels be used in the assessment and evaluation of student performance.

Results interpretation panels, which would identify expectations of the public, could be particularly useful when benchmarks or norms have not been established; that is, when the tests are new or have been substantially revised. The panels would also be appropriate in other instances because norms alone may not provide enough

information for making judgements about how well the students have performed. Some of the difficulties which MACOSA's results interpretation panels encountered made it clear that, when panels are used, their judgements must be based on a firm understanding of the program being assessed and the techniques being used in the assessment. This could be achieved through early and appropriate involvement of the panels in the particular assessment program.

9. THAT a computerized item banking project, including study of inservice and consultation components, be commissioned in order to study the feasibility of this approach in serving a number of purposes for groups with responsibilities for assessment at the provincial, school system and classroom levels.

Items for such a data bank would be developed, validated and selected both locally and provincially and filed centrally. The items would be based upon a large number of objectives. Input from the school system level would provide for continuous additions of items to the bank and continuous validation of items. These items could be combined to make common tests for various agencies (ranging from Alberta Education to individual teachers) and tests tailor-made to specific purposes such as initial placement, diagnosis of learning handicaps and determining the effectiveness of instructional programs.

In February 1979 MACOSA received information on the establishment of the Ontario Assessment Instrument Pool, of which procedures for collecting, retrieving, using, updating and replacing items form a significant part. MACOSA advises that Alberta Education should keep up to date on the progress of this activity.

10. THAT the Regional Offices of Alberta Education provide school jurisdictions with consultative services regarding student assessment programs.

Under MACOSA's proposals most of the responsibility for student assessment remains at the local level; however, findings from the *Alberta Grade 12 Examination Study*, particularly the observations of school superintendents, suggest that local expertise in student assessment varies among school jurisdictions. The regional offices are an appropriate agency for providing support services regarding evaluation techniques. These services could assist in at least three areas: improving instruction, developing expertise of teachers in evaluation, and providing some degree of comparability both in evaluation techniques and the proportion of high and low marks awarded within and among school systems.

II. The Departmental Examination Question

In October 1977 MACOSA reviewed the findings of the *Alberta Grade 12 Examination Study* and concluded that: a) serious concerns had been expressed about educational quality; b) there was a general desire that action be taken to ensure or enhance this quality; c) there was considerable reluctance on the part of the various

publics to return to the pre-1973 mandatory type of departmental examinations; d) although the system of accreditation of local school systems to award final marks in what had been "examinations subjects" was not working as well as it might, the system could be made to work. Subsequently MACOSA made an interim general recommendation to the Minister that a final decision on grade 12 departmental examinations be deferred for at least three years and that "the necessary steps be taken to enhance the effectiveness of the existing policy with respect to student evaluation". MACOSA did not at that time specify what constituted the "necessary steps", but did indicate that advice on this issue would be offered after completion of several other MACOSA studies.

In November 1977 the Minister announced to the Legislative Assembly that in view of the evidence of variation in school system evaluation policies and practices, Alberta Education would "intensify supervision of school evaluation policies and procedures to reduce major discrepancies among systems and schools". Additionally, the Curriculum Branch would "continue to more precisely specify learning objectives in existing programs of study and curriculum guides for the information of students, teachers, and the public", a process which had commenced in response to a recommendation of the Curriculum Policies Board at about the same time as the establishment of MACOSA. The Minister noted also that the question of mandatory grade 12 departmental examinations would be reviewed after MACOSA filed its final report.

Pursuant to these statements from the Minister, Alberta Education in cooperation with representatives of the teachers', trustees', and superintendents' organizations developed a statement on *Guidelines for Developing Student Evaluation Policy in Alberta School Systems* (1978). These guidelines provided a model for local jurisdictions. In the fall of 1978 the guidelines were distributed to all school systems and schools. In addition, Alberta Education continues to analyze high school marks awarded by school systems and make inquiries into cases of unexplained discrepancies in the proportions of high and low marks assigned within particular school systems and schools. Since 1977 all school systems which had been without local policies on student evaluation have adopted such policies. Consequently, awareness of the requirement for the use of fair, and to some degree comparable, means for evaluation has increased. Work continues on securing the effectiveness of system policies at the school level.

All the foregoing developments provide background for the following recommendation.

The Minister's Advisory Committee on Student Achievement recommends:

11. THAT mandatory grade 12 departmental examinations for the purpose of awarding final marks not be reinstituted.

As a result of weighing the evidence received from MACOSA's various studies and advisory groups, MACOSA concludes that the principal responsibility for evaluation of student achievement for purposes of promotion through the several grade levels and for the awarding of diplomas should continue to reside with local school systems. Alberta Education should continue to provide assistance to improve practices in student assessment and should regularly monitor levels of achievement.

The argument against reinstitution of mandatory departmental examinations in their historical form is based on two main contentions: first, that these examinations have not provided their usually assumed standard of evaluation; and second, that teacher-assigned marks provide as good or better predictions of later academic success in post-secondary studies than do marks from common examinations.

On the first point, the *Alberta Grade 12 Examination Study* discussed serious problems and disadvantages associated with the departmental examinations system terminated in 1973. This study noted that a number of popular beliefs concerning the benefits of the examinations, on which success or failure of students in matriculation subjects was based, proved to be illusory if not myth or false hopes. Contrary to popular belief that the examinations maintained educational standards from one year to the next, the examinations actually did not serve this purpose. Changes in the content of courses, changes in the emphases given (the tests stressed different sections of a course from year to year), and the need to develop a completely new set of questions for each test date all prevented the tests from being sufficiently comparable to ensure a viable basis from which to establish an enduring standard of achievement. A statistical procedure was adopted to fix the marks to result in approximately the same proportion of As, Bs, Cs, Ds and failures from year to year. Beginning in 1966, the distribution of marks in each subject was fixed to parallel the distribution of scores which students enrolled in the particular subject attained on a general ability test — the *School and College Ability Test* (SCAT). If 8 percent of the students in the subject obtained A-range scores on the SCAT, then 8 percent of the marks given on the departmental examination in that subject would be assigned to the A-range. Since the difficulty of the tests varied, the only "standard" involved from year to year came about from the system used for assigning the relative proportions of marks at different grade levels. The only validity of the examinations for purposes of comparison was how well a particular student ranked in comparison with classmates writing the same examination at the same time. Comparisons between individuals or groups over time were invalid because of differences in the examinations used each time a subject was tested.

Consequently, the examinations never provided a measure of year-to-year trends for the purpose of monitoring achievement standards; there was merely a prevalent but erroneous belief that provincial trends were being tracked over time. The best the examinations system could do, given the nature of the tests and the scoring

system, was to set comparable standards of achievement from school to school within a single semester.[21]

On the other hand, MACOSA's proposal for periodic testing to determine levels of achievement on a provincial basis will provide the information which many groups want regarding general levels of achievement not only in a given year but over longer periods as well.[22] MACOSA's proposal meets the demand for maintenance of standards in a way that would not be possible through reinstitution of departmental examinations in their historical form.

Regarding the second contention — the question of the relative efficiency of grade 12 departmental examinations versus teacher-assigned marks as predictors of academic success in post-secondary education — little evidence is available. MACOSA found that post-secondary educational institutions vary greatly regarding marking procedures, types of grades assigned, and methods of storing information on marks. Given the time line for MACOSA's reports, and the mandate of the Committee, MACOSA had no opportunity to systematically study data on post-secondary marking patterns. However, the one institution-wide study which came to MACOSA's attention (Mehra, 1977) indicates that at the University of Alberta, which is the largest post-secondary educational institution in the province, teacher-assigned marks in university entrance subjects are generally slightly better as predictors of success in university programs than were the departmental examination marks assigned before the 1973 change to the accreditation system. Thus the available evidence confutes popular belief and does not support a return to the departmental examinations in order to improve the screening of applicants for entrance to at least the university sector of post-secondary educational programs.

There are a number of speculations, generally validated through experience, that support the view that teacher-assigned marks are more likely to reflect the student's true level of achievement than does the score on a single examination. Teacher marks presumably would be based on students' demonstrated competence and achievement in a variety of assignments over the full term of a course. The marks should reflect a diversity of measurements and observations of attitudes and accomplishments ranging over the full spectrum of educational goals, some of which are difficult or impossible to test in the standard examination setting. Further, teacher-assigned marks should minimize the influencing of results of a machine-scored test through cramming and "test wiseness" and the effects of particularly adverse psychological conditions which some students might be endur-

[21] For a fuller review of the history of provincial examination practices in Alberta, see *Roles in Student Evaluation and Research* (Alberta Education, 1976), pp. 2-6. This publication is available on request from the Communications Branch of Alberta Education.

[22] See Recommendation #1, p. 66, which contains MACOSA's proposal for periodic testing at the grades 3, 6, 9 and 12 levels.

ing during the time period set aside for a specific examination. The impact of all the aforementioned influences provides further reasons for rejecting the historical type of departmental examination as a basis for awarding final marks.

MACOSA's proposals for a testing program and consultative services[23] promise a broader and sounder approach to maintaining standards than reversion to the departmental examinations system could. MACOSA's proposals are designed to provide standards and convey information to interested groups while allowing the promotion of students and assignment of final marks to remain predominantly local matters except for certain types of appeal, as mentioned in Recommendation #13 following. Since MACOSA's tests would remain valid for the curriculum from year to year (until the curriculum is revised substantially), and since the tests would be of equal difficulty from year to year, the tests would provide an indication of standards over time — a desirable role which many people mistakenly thought was filled by the mandatory grade 12 departmental examinations.

12. THAT grade 12 departmental examinations continue to be made available in January and June for school systems that wish to use them.

Evaluation of students for purposes of promotion should continue to be primarily the responsibility of local school jurisdictions. The departmental examinations should remain available, however, for those jurisdictions not wishing to exercise this responsibility.

13. THAT Alberta Education continue to provide grade 12 appeal examinations for their current uses.

The appeal examinations are still useful for special cases involving individual students, such as those requiring external criteria for scholarships applications, candidates for adult diplomas, and students who wish to challenge teacher-assigned marks.

III. Proposed Student Evaluation Policies Board and Referrals

The Minister's Advisory Committee on Student Achievement recommends:

14. THAT Alberta Education establish a Student Evaluation Policies Board composed of representatives of stakeholder groups to advise the Minister of Education on matters relative to student achievement.

MACOSA considers that continuing public input is vital to effective and responsive policy decisions and actions on provincial student evaluation, and therefore that a structure which would serve in such an advisory capacity should be established.

The terms of reference for the proposed Student Evaluation Policies Board should include the following:

[23] See Recommendation #10, p. 69.

Function To advise the Minister of Education on matters relative to student achievement.

Membership

1. The Board shall consist of a chairman and nine appointed members.
2. The Minister shall appoint, for a term of up to three years, members from lists of nominees submitted by each of the following groups or organizations:
   a. Alberta Chamber of Commerce (one member)
   b. Alberta Federation of Home and School Associations (one member)
   c. Alberta Federation of Labour (one member)
   d. Alberta School Trustees' Association (one member)
   e. Alberta Teachers' Association: one representative for the elementary level and one representative for the secondary level (two members)
   f. Conference of Alberta School Superintendents (one member)
   g. Committee on the Articulation of High Schools and Post-Secondary Educational Institutions: representatives of the post-secondary educational community (two members)
3. The chairman of the Board shall be an Associate Deputy Minister of Education.
4. The Minister shall appoint from Alberta Education a non-voting Executive Secretary and such administrative assistance as required. This administrative support may be assigned from one or more branches of Alberta Education.

Term of Office The Board shall sit at the pleasure of the Minister.

Specific Terms of Reference

1. The Board shall advise the Minister with respect to provincial policies related to the assessment of student achievement in Alberta.
2. The Board may advise the Minister relative to appropriate evaluation guidelines at the school system level.
3. The Board shall advise the Minister from time to time as to the adequacy of student achievement in specific curricular programs.
4. The Board may recommend the establishment of committees to provide direction to any studies or projects that the Minister may commission on subjects relating to student achievement. Terms of reference for such committees are to be established by the Board.
5. The Board shall attend to such other matters relating to the evaluation of student achievement as may be referred to it by the Minister.

The proposed Student Evaluation Policies Board should keep in close contact with the Curriculum Policies Board because the work of each board has implications for the other.

15. THAT recommendations submitted by the steering committees and/or researchers for the following MACOSA test development studies be referred to the Student Evaluation Policies Board and/or to Alberta Education:
    a. Listening and Speaking
    b. Attitudes Towards the World of Work
    c. Attitudes Towards School Subjects
    d. Assessment of Affective Objectives in the Social Studies

In Recommendation #3, MACOSA dealt specifically with the recommendations regarding revision and norming of the tests developed for studies a-c above. Recommendation #5 stems from the report on *Assessing Affective Objectives in the Social Studies*. Because of the limited time available for discussing the other recommendations arising out of the test development studies, MACOSA recommends that those recommendations be referred for future consideration.

16. THAT MACOSA's summary of research findings and conclusions regarding factors affecting student achievement be referred to the Student Evaluation Policies Board and/or Alberta Education for further study and appropriate action.

The MACOSA achievement studies revealed certain strengths and weaknesses in student performances but because of a lack of information the researchers did not attempt to explain the levels of performance. In this report MACOSA has speculated in a few instances, keeping in mind that further study of educational and environmental factors affecting student achievement (as described in Chapter 4) is needed before any causes can be definitively identified.

17. THAT MACOSA's conclusions and comments regarding levels of achievement be referred to the Student Evaluation Policies Board, Alberta Education and/or other concerned agencies for further study and appropriate action.

As noted in the rationale for Recommendation #16, reasons for high and low achievement levels are not known. Because of the lack of such information and the limited time frame, MACOSA could not state specifically what actions should be taken to improve student achievement. Weaknesses and strengths could be related to the appropriateness of the curriculum, implementation of curriculum in the classroom, the characteristics of students, or as indicated in Chapter 4, interactions

among curriculum, instruction and student traits, or numerous other factors. It is MACOSA's hope that the findings and conclusions of the achievement studies will be considered carefully by agencies and boards making decisions about curriculum, instruction, and evaluation, and by those who are doing research on student achievement.

GLOSSARY

This glossary provides non-technical definitions of some standard terms used in studies on student achievement.

ACCREDITATION In Alberta, the Department of Education's delegation of its role in evaluating matriculation subjects to the schools, with the proviso that the schools act within stated guidelines. The guidelines established in 1973 included following prescribed curricula, adhering to minimum time allotments for certain school subjects, and hiring certificated teachers.

ACHIEVEMENT TEST A test designed to measure a students' grasp of some body of knowledge or his proficiency in certain skills, usually as a result of specific instruction.

AFFECTIVE Pertaining to the feelings, emotions, or attitudes. Affective outcomes of education involve feelings more than understandings. A person's likes and dislikes, ideals and values are some of the affective outcomes that education may develop.

ATTITUDE TEST A test designed to measure the tendency to respond favorably or unfavorably towards specific groups, institutions or objects.

CLUSTER OF ITEMS A group of test questions related to the same content area or the same program objective.

COGNITIVE Having to do with perception, knowledge and understanding. Cognitive outcomes of education involve learning to know or understand the facts or information included in the subject of study. *(See* Affective.)

CONTENT VALIDITY (OF AN EDUCATIONAL ACHIEVEMENT TEST): The extent to which the items in the test adequately sample the areas of subject matter and the skills that a course of instruction has aimed to teach.

CORRELATION The extent to which two variables are related; a correlation may have either a positive or a negative value. (*See* p. 19 for examples.)

CRITERION-REFERENCED TEST Any test designed to measure level of attainment in specific content or skill areas against a predetermined standard. (*See* Norm-referenced test.)

CURRICULUM 1) a group of courses and planned experiences which a student has under the guidance of a school, 2) a general overall plan of the content or specific materials of instruction that the school should offer the student by way of qualifying him for graduation or certification.

ITEM In test construction, a simple testing task (test question) which provides a means to measure achievement of a knowledge, skill or attitudinal objective.

MATRIX SAMPLING A sampling technique in which both students and test items are sampled. A test instrument consists of several different sets of questions. Each student sampled responds to only one set of questions, which is randomly

assigned to him. This method is useful for evaluating groups without asking any one student to answer too many questions. (*See* pp. 9, 33.)

MEAN A measure of the "arithmetic average" which is found by summing all values and dividing by the number of cases.

MEDIAN The middle point in a series of measurements after all values have been placed in sequence.

NORM 1) average or typical test score, or other measure, for members of a specific group, 2) standard or criterion for judgement.

NORM-REFERENCED TEST A test designed to measure how well a student does in relation to the group.

PILOT TESTING Trial testing on a small scale of a new test in order to see if it is workable and warrants wider adoption.

RANGE The difference between the highest and lowest scores in a distribution; used as a measure of variability or spread of scores. For example, the range between 42 and 85 is 43 points.

RAW SCORE The number of points an individual receives on a test.

RELIABILITY The degree to which a test measures accurately and consistently (sometimes referred to as test-retest reliability). For example, a reliable test will produce similar results when administered to the same group at different times.

SEMANTIC DIFFERENTIAL TECHNIQUE A method of testing in which respondents are asked to indicated by placing a mark on a rating scale, the meaning they attach to a concept or the attitude they hold. (*See* p. 56 of this report for a sample item.)

STAKEHOLDER Person or group concerned about issues in basic education such as students, parents, teachers, administrators, school boards, employers, post-secondary educational institutions and the general public.

STANDARD DEVIATION A measure of the spread of scores around the mean or average score.

STANDARDIZED TEST As opposed to a teacher-made test, a standardized test is designed to be administered and scored under uniform testing conditions, and norms are usually provided.

VALIDITY The extent to which a test measures what it purports to measure or does the job for which it is intended. (*See* also Content validity.)

BIBLIOGRAPHY

Alberta Education. *Goals of basic education* (Int. ed.), 1975.

Alberta Education. *The goals of basic education for Alberta*, 1978.

Alberta Education. *Roles in student evaluation and research*, 1976.

Alberta Education. *Guidelines for developing student evaluation policy in Alberta school systems*, 1978.

Bevan, B., Gajadharsingh, J., & Gommeringer, B. *Ways and means: evaluation of communication skills*. Alberta Education, 1977.

Canadian Institute for Research in the Behavioral and Social Sciences. *Alberta language arts (reading and writing) achievement study*. Alberta Education, 1979.

Carmichael, G., & Rees, R. E. A survey of reading achievement in Alberta schools. *Alberta Journal of Educational Research*, 1955, 1 (1), 18-33.

Chandler, B. et al. *Feasibility study on assessing student achievement in social studies*. Alberta Education, 1977.

Clarke, S. C. T., & Nyberg, V. (University of Alberta). *Development of school subjects attitude scales*. Alberta Education, 1979.

Clarke, S. C. T., & Nyberg, V., & Worth, W. (Alberta Advisory Committee for Educational Studies). *Alberta grade 3 achievement study*. Alberta Education, 1978.

Clarke, S. C. T., Nyberg, V., & Worth, W. H. (Alberta Advisory Committee for Educational Studies). *Edmonton grade 3 achievement study: 1956-1977 comparisons*. Alberta Education, 1977.

Coleman, J. et al. *Equality of educational opportunity*. Washington, D. C.: U.S. Document Printing Office, 1966.

Dumont, F. J. *Alberta grade 12 examination study*. Alberta Education, 1977.

Ebel, R. L. *Essentials of education measurement*. Englewood Cliffs, N.J.: Prentice-Hall, 1972.

Edmonton Public School District. *Listening and speaking objectives*. Alberta Education, 1979.

Fedigan, L. *School-based elements related to achievement: a review of the literature*. Alberta Education, 1979.

Gay, G. *Elements related to student success in schooling and education: a review of the literature*. Alberta Education, 1979.

Gommeringer, B., McFetridge, P., & Plattor, E. *Ways and means: evaluation of listening and speaking skills*. Alberta Education, 1979.

Good, C. V. (Ed.) *Dictionary of education*. New York: McGraw-Hill, 1973.

Lopatka, R. *Report on the development and administration of the high school social studies achievement test*. Alberta Education, 1978.

Lopatka, R. *Development of models for assessing affective objectives in the social studies*. Alberta Education, 1979.

Maguire, T., & Romaniuk, E. (University of Alberta). *Development of scales on attitudes towards the world of work*. Alberta Education, 1979.

Marchak, N. et al. (Edmonton Public School District). *Assessing communication skills: a review of the literature*. Alberta Education, 1979.

Mehra, N. *High school grades and university performance*. University of Alberta, 1977.

Olson, A., Sawada, D. & Sigurdson, S. (University of Alberta). *Alberta assessment of school mathematics*. Alberta Education, 1979.

Plattor, E., Unruh, W., & Muir, L. (University of Calgary). *Test development for assessing achievement in listening and speaking*. Alberta Education, 1979.

Rosenthal, R., & Jacobson, L. *Pygmalion in the classroom: teacher expectations and pupils' intellectual development*. New York: Holt, Rinehart and Winston, 1968.

Rusnack, T. A. Evaluating education in Alberta schools. Unpublished manuscript, Alberta Education, 1975.

Rusnack, T. A. Accreditation in Alberta education. Unpublished doctoral dissertation, University of Calgary, 1977.

Sax, G. *Principles of education measurement and evaluation*. Belmont, Calif.: Wadsworth, 1974.

Stringham, B. L. *Next steps in accreditation*. Alberta Education, 1975.

Treasure, M. *Feasibility study in assessing student achievement in science*. Alberta Education, 1977.

Treasure, M. *Alberta science achievement study*. Alberta Education, 1979.

Warren, R. *A study re high school graduation requirements*. Alberta Education, 1975.

APPENDIX A

Ministerial Order Establishing the Minister's Advisory Committee on Student Achievement

Pursuant to section 6 of the Department of Education Act, I, Julian Koziak, Minister of Education, hereby establish the Minister's Advisory Committee on Student Achievement in accordance with the attached Schedule.

Schedule

1. **Function** The Committee shall advise the Minister of Education relative to the aspects of student achievement noted in section 4. Recommendations and reports will be forwarded for consideration by the Minister when available and at or near the end of the Committee's term of office.

2. **Membership**

   2.1 The Minister shall appoint a Committee consisting of nine members.

   2.2 Each of the following shall nominate one representative for the Minister's consideration for appointment:

   (a) Alberta Chamber of Commerce

   (b) Alberta Federation of Home and School Associations

   (c) Alberta Federation of Labour

   (d) Alberta School Trustees' Association

   (e) Alberta Teachers' Association

   (f) Alberta Department of Advanced Education and Manpower

   (g) Committee on the Articulation of High School and Post-secondary Educational Institutions

   (h) Conference of Alberta School Superintendents

   2.3 The Chairman of the Committee shall be appointed by the Minister from the Alberta Department of Education.

   2.4 The Minister shall appoint from the Alberta Department of Education a non-voting Executive Secretary and Project Director.

3. **Term of Office** The Committee shall sit from November 1, 1976 to December 31, 1978.*

4. **Terms of Reference**

   4.1 The committee shall study the elements of student achievement and will select appropriate elements as a focus or guide to its activities.

   4.2 Aspects of student achievement which shall be studied include the following:

   (i) Rationale — the purposes, basic principles, assumptions underlying student achievement

*On December 4, 1978 the concluding date was changed to March 31, 1979.

(ii) Definitions of student achievement

(iii) Scope — instruction and achievement

(iv) Standards and their maintenance

4.3 Committee recommendations shall attend to the following matters:

(i) Commissioning of papers, studies and/or projects concerning the dimensions of student achievement identified above, provided however that each such proposal which will result in an expenditure of more than $3000 shall be submitted to the Minister for his approval before its commissioning

(ii) Considering the effect of the non-compulsory nature of Grade XII Departmental Examinations on the quality of education in Alberta

(iii) Ways and means to assess, maintain and improve student achievement (based on findings from (i) and other considerations)

(iv) Such other matters as may be referred to the Committee by the Minister

4.4 It shall be the responsibility of the Minister's Advisory Committee to recommend the establishment of steering committees to provide direction to the studies and projects commissioned by the Minister relating to student achievement. Terms of reference for such steering committees are to be set out by the Advisory Committee. Specific responsibilities for steering committees will include:

(a) Developing study designs or requests for proposals in preparation for establishing a task force activity or commissioning a study

(b) Ensuring that progress reporting and phase completion deadlines as provided for are being adhered to as closely as possible

(c) Ensuring that the parties to a study are well informed as to its progress on an on-going basis and where appropriate, recommending public information releases

(d) Providing whatever technical or policy advisory assistance to the investigator(s) as may be required from time to time

(e) When investigations are conducted in stages, making recommendations to the Advisory Committee relative to the following:

(i) as to whether or not the study should proceed into the next stage

(ii) as to the sufficiency of funds allocated in respect of subsequent research stages, where progress into the next stage is recommended, within the total contract limit

(iii) whether further work should be done on the stage just completed or

(iv) whether a study or project should be terminated

(v) upon completion of the final stage of a project, as to disposition of the final report

(f) With regard to each commissioned study, assuring that as many copies of the final report together with an executive summary as may have been provided for in any contract are tabled with the Advisory Committee upon completion

(g) Such other matters as may be referred to a steering committee by the Committee.

APPENDIX B — The Goals of Basic Education for Alberta

Introduction

Goals are statements which indicate what is to be achieved or worked toward. In relation to basic education, goals serve several functions:

(1) they identify the distinctive role of the school and its contribution to the total education of youth;

(2) they provide purpose and direction to curriculum planning, implementation and evaluation;

(3) they enable parents, teachers and the community at large to develop a common understanding of what the schools are trying to achieve.

Society must periodically re-examine the goals of its schools. Changes in emphasis and minor adjustment of the basic goals may be required from time to time to keep pace with social change.

This statement of goals is to direct education for grades 1 through 12 in Alberta schools. It is the basis from which specific objectives for various subjects and grades shall be developed.

While the school makes a very important contribution to education, it is only one of the agencies involved in the education of youth. The home, the church, the media and community organizations are very significant influences on children. It is useful, therefore, to delimit the role of schooling in education. Education refers to all the learning experiences the individual has in interacting with the physical and social environment; it is a continuing and lifelong process. Schooling, which has a more limited purpose, refers to the learning activities planned and conducted by a formally structured agency which influences individuals during a specified period. There is, of course, a very close relationship between schooling and education — the learning which occurs in school influences and is influenced by what is learned outside the school.

Goals of Schooling

Schooling, as part of education, accepts primary and distinctive responsibility for specific goals basic to the broader goals of education. Programs and activities shall be planned, taught and evaluated on the basis of these specific goals in order that students:

— Develop competencies in reading, writing, speaking, listening and viewing.

— Acquire basic knowledge and develop skills and attitudes in mathematics, the practical and fine arts, the sciences, and the social studies (including history and geography), with appropriate local, national and international emphasis in each.

— Develop the learning skills of finding, organizing, analyzing and applying information in a constructive and objective manner.

— Acquire knowledge and develop skills, attitudes and habits which contribute to physical, mental and social well-being.

— Develop an understanding of the meaning, responsibilities and benefits of active citizenship at the local, national and international levels.

— Acquire knowledge and develop skills, attitudes and habits required to respond to the opportunities and expectations of the world of work.

Because the above goals are highly interrelated, each complementing and reinforcing the others, priority ranking among them is not suggested. It is recognized that in sequencing learning activities for students some goals are emphasized earlier than others; however, in relation to the total years of schooling, they are of equal importance.

In working toward the attainment of its goals, the school will strive for excellence. However, the degree of individual achievement also depends on student ability and motivation as well as support from the home. Completion of diploma requirements is expected to provide the graduate with basic preparation for lifelong learning. Dependent on program choices, the diploma also enables job entry or further formal study.

Goals of Education

Achievement of the broader goals of education must be viewed as a shared responsibility of the community. Maximum learning occurs when the efforts and expectations of various agencies affecting children complement each other. Recognizing the learning that has or has not occurred through various community influences, among which the home is most important, the school will strive to:

— Develop intellectual curiosity and a desire for lifelong learning.

— Develop the ability to get along with people of varying backgrounds, beliefs and lifestyles.

— Develop a sense of community responsibility which embraces respect for law and authority, public and private property, and the rights of others.

— Develop self-discipline, self-understanding and a positive self-concept through realistic appraisal of one's capabilities and limitations.

— Develop an appreciation for tradition and the ability to understand and respond to change as it occurs in personal life and in society.

— Develop skills for effective utilization of financial resources and leisure time and for constructive involvement in community endeavors.

— Develop an appreciation for the role of the family in society.

— Develop an interest in cultural and recreational pursuits.

— Develop a commitment to the careful use of natural resources and to the preservation and improvement of the physical environment.

— Develop a sense of purpose in life and ethical or spiritual values which respect the worth of the individual, justice, fair play and fundamental rights, responsibilities and freedoms.

The ultimate aim of education is to develop the abilities of the individual in order that he might fulfill his personal aspirations while making a positive contribution to society.

Adopted by the Legislative Assembly of
The Province of Alberta
May 15, 1978

APPENDIX C — Sources and Availability of MACOSA Studies

Executive summaries or condensed versions of each major MACOSA study are available on request while supplies last from the Planning and Research Branch, Alberta Education, 9th Floor, Devonian Building, 11160 Jasper Avenue, Edmonton, Alberta T5K 0L2; Ph. (403) 427-8217.

The full reports of each study are available on loan from 7 Regional Offices of Alberta Education and from the Faculty of Education libraries of many Canadian universities.

The reports may be purchased from the ERIC Document Service (on microfiche or hard copy), P.O. Box 190, Arlington, Va. 22210, or from Micro-Media Limited (in full-sized paper reprint form). Micro-Media's address is Box 502, Station S, Toronto, Ontario, Canada M5M 4L8.

APPENDIX D — MACOSA Administrative and Consultative Staff

Dr. Harry Sherk, Executive Secretary and Project Director (Alberta Education, Planning and Research Branch)

Mr. Ray LaFleur, Administrative Secretary (Alberta Education, Communications Branch)

Dr. Wesley Eddy (Alberta Education, Planning and Research Branch)

Dr. Jack Reid (Alberta Education, Student Evaluation and Data Processing Services Branch)

Dr. Clarence Rhodes (Alberta Education, Planning and Research Branch)

Dr. Morris Treasure (Alberta Education, [illegible] Research Branch)

Mr. Bruce V. Decoux, Administrative [illegible] (Assistant to Associate Deputy Minister of Alberta Education)

Ms. Anita Jenkins, Editor/Writer (Alberta Education)

Mr. Lalta Kunjbehari, Administrative Intern (Alberta Education, Planning and Research Branch)

APPENDIX E

MACOSA Results Interpretation Panels

The following persons (members of the general public, parents, educators and representatives of other interest groups) served on the MACOSA results interpretation panels, September 13-15, 1978 at the Mayfield Inn in Edmonton.

Elementary Language Arts Panel

Mr. J. Rennie (Chairman), Language Arts Consultant, Alberta Education

Ms. S. Capowski, Language Arts Consultant, Edmonton Public School Board

Ms. I. Evans, Education Committee Member, Strathcona County

Ms. J. Koper, Area Superintendent, Calgary Board of Education

Dr. W. Laing, Faculty of Education, University of Alberta

Mrs. L. Milner, Curriculum Policies Board Member

Mr. A. Schaufert, Supervisor, Red Deer Public School District

Mrs. E. Townend, President, Alberta School Trustees' Association

Dr. B. Honert (Recorder), Language Arts Consultant, Alberta Education

Secondary Language Arts Panel

Dr. M. Thornton (Chairman), Associate Director of Curriculum, Alberta Education

Mrs. I. Boston, Communications Officer, Banff Centre

Mr. T. Cashman, Public Relations Director, Alberta Government Telephones

Ms. M. Layman, Director of Public Affairs, Alberta Culture

Ms. J. Lubin, Teacher, Edmonton Public School Board

Dr. G. Martin, Faculty of Education, University of Alberta

Mrs. J. Puckett, Canadian Union of Public Employees, Lethbridge

Mr. J. Ronjom, Superintendent of Schools, Rocky Mountain School Division

Mr. A. Stephenson, Director of Academic Upgrading, Fairview College

Mr. D. Stewart, Teacher, Strathcona County

Ms. A. Jenkins (Recorder), MACOSA Editor/Writer, Alberta Education

Elementary Mathematics Panel

Mr. A. Anderson (Chairman), Mathematics Consultant, Alberta Education

Mrs. J. Anderson, President, Alberta Federation of Home and School Associations

Dr. G. Cathcart, Faculty of Education, University of Alberta

Ms. F. Craigie, Trustee, Red Deer Public School District and Curriculum Policies Board Member

Mr. R. Holt, Teacher, Edmonton Public School Board

Mr. B. Stonell, Mathematics Consultant, Alberta Education

Mr. V. Taylor, Training Officer, Syncrude Canada

Mr. E. Warnica, Mathematics-Science Coordinator, Lethbridge Public School District

Mr. A. Peddicord (Recorder), Mathematics Consultant, Alberta Education

Secondary Mathematics Panel

Mr. R. Daly (Chairman), Mathematics Consultant, Alberta Education

Mrs. B. Ballhorn, Alberta Federation of Home and School Associations, Wetaskiwin

Mr. M. Bye, Supervisor, Calgary Board of Education

Mr. D. Fisher, Teacher, Calgary Board of Education

Mrs. D. Freeborn, Physiotherapist, Fort Saskatchewan

Dr. A. Gibb, Faculty of Education, University of Calgary

Dr. D. MacNeill, Public Service Commission, Government of Alberta

Mr. J. Percevault, Faculty of Education, University of Lethbridge

Mrs. J. Rubin, Housewife, Lethbridge

Mr. G. Wooldridge, Chartered Accountant, Willetts MacMahon and Company, Edmonton

Mr. W. Lencucha (Recorder), Mathematics Consultant, Alberta Education

Elementary Science Panel

Mr. G. Popowich (Chairman), Science Consultant, Alberta Education

Mrs. D. Christie, Alberta Federation of Home and School Associations and Curriculum Policies Board Member

Mr. B. Galbraith, Supervisor, Edmonton Public School Board

Mr. A. Jamha, Education Committee Member, Strathcona County

Mrs. J.Lore, Past-President, Canadian Federation of Home and School and Parent-Teacher Federations

Dr. J. Marean, Faculty of Education, University of Calgary

Dr. J. Thiessen (Recorder), Consultant, Planning and Research Branch, Alberta Education

Secondary Science Panel

Dr. G. Gay (Chairman), Science Consultant, Alberta Education

Dr. F. Bentley, Faculty of Agriculture, University of Alberta

Mr. W. Brooks, Executive Assistant, Professional Development, Alberta Teachers' Association

Mr. M. Lynch, Science Supervisor, Edmonton Catholic Schools

Mr. H. Marcellus, Teacher, Edmonton Public School Board

Dr. M. Sillito, Retired; Former Coordinator of Professional Development, Alberta Teachers' Association

Mr. T. Styra, Retired teacher, Radway, Alberta

Mrs. N. Vester, Family Living Specialist, Home Economics Branch, Alberta Education

Mrs. J. Westerlund, Alberta Federation of Home and School Associations, Oyen, Alberta

Dr. A. Zelmer, Dean of Nursing, University of Alberta

Mr. F. Nordby (Recorder), Science Consultant, Alberta Education

Secondary Social Studies Panel

Mr. R. Johnson (Chairman), Social Studies Consultant, Alberta Education

Mr. R. Baker, Principal, Edmonton Public School Board

Dr. K. Dueck, Faculty of Education, University of Calgary

Mr. P. Gibeau, Family Education Officer, Alberta Consumer and Corporate Affairs

Mrs. D. McCullough, Retired, Former Member of University of Alberta Senate

Mr. Wm. McIntosh, Supervisor, Calgary Board of Education

Ms. S. Merchant, Consumer and Corporate Affairs Canada

Mr. W. Romanko, Assistant Director of Continuing Education, Northern Alberta Institute of Technology

Mr. G. Schuler, Rural Education and Development Association, Edmonton

Mr. R. Webb, Department Head, Edmonton Public School Board

Mr. K. Mergle (Recorder), Teacher, Alberta Correspondence School

APPENDIX F — MACOSA Steering Committee Members

Alberta Grade 3 Achievement Steering Committee

Dr. C. Rhodes (Chairman), Consultant, Alberta Education

Mr. J. Anderson, Director, Education Services, Alberta School Trustees' Association

Mr. W. Brooks, Executive Assistant, Professional Development, Alberta Teachers' Association

Mr. A. Myhre, President, Conference of Alberta School Superintendents

Language Arts (Reading and Writing) Achievement Steering Committee

Mr. J. Wood (Chairman), Associate Director, Alberta Education

Dr. C. Braun, Faculty of Education, University of Calgary

Mrs. I. Hargreaves, Language Arts Supervisor, Edmonton Public School Board

Mr. J. Rennie, Language Arts Consultant, Alberta Education

Dr. C. Rhodes (Recorder), Consultant, Alberta Education

Dr. M. Thornton (Consultant to the Committee), Associate Director, Alberta Education

Mathematics Achievement Steering Committee

Dr. C. Rhodes (Chairman), Consultant, Alberta Education

Mr. A. Anderson, Mathematics Consultant, Alberta Education

Dr. A. Gibb, Faculty of Education, University of Calgary

Mr. E. Warnica, Mathematics-Science Coordinator, Lethbridge Public School District

Dr. M. Treasure (Recorder and Consultant to the Committee), Alberta Education

Science Achievement Steering Committee

Dr. L. Tolman (Chairman), Associate Director, Alberta Education

Dr. Wm. Holliday, Faculty of Education, University of Calgary

Mr. I. Ibuki, Assistant Superintendent of Schools, County of Lethbridge

Mr. M. Lynch, Science Supervisor, Edmonton Catholic Schools

Dr. C. Rhodes (Recorder), Consultant, Alberta Education

Dr. M. Treasure (Consultant to the Committee), Consultant, Alberta Education

Social Studies Achievement Steering Committee

Mr. R. Johnson (Chairman), Social Studies Consultant, Alberta Education

Mr. F. Bowen, Teacher, Lethbridge Public School District

Mr. Wm. McIntosh, Social Studies Supervisor, Calgary Board of Education

Dr. M. Van Manen, Faculty of Education, University of Alberta

Grade 12 Examination Study Steering Committee

Mr. L. Tymko (Chairman), Alberta School Trustees' Association

Mrs. M. Fitch, Homemaker, Calgary

Mrs. S. Forbes, Trustee, Edmonton Public School Board

Mr. J. Layton, Teacher, Innisfail High School
(Mr. W. Brooks of the Alberta Teachers' Association substituted for Mr. Layton on some occasions.)

Dr. T. Maguire, Faculty of Education, University of Alberta

Dr. K. Sauer, Superintendent, Medicine Hat Public School Board

Mr. R. LaFleur (Secretary), Assistant to the Director of Information Services, Alberta Education

Listening and Speaking Study Steering Committee

Mr. B. Gommeringer (Chairman), Language Arts Consultant, Alberta Education

Dr. W. Laing, Faculty of Education, University of Alberta

Dr. Wm. Washburn, Calgary Board of Education

Dr. M. Thornton (Consultant to the Committee), Associate Director, Alberta Education

Dr. C. Rhodes, Consultant, Alberta Education (Recorder and Consultant to the Committee)

Attitude Tests Steering Committee

Dr. T. Mott (Chairman), Counselling and Guidance Supervisor, Alberta Education

Dr. T. Maguire, Faculty of Education, University of Alberta
(Dr. W. Unruh, Faculty of Education, University of Calgary subsequently replaced Dr. Maguire.)

Dr. H. Mosychuk, Assistant Superintendent of Research and Evaluation, Edmonton Public School Board

Mr. T. Tomko, School Psychologist, Edmonton Catholic Schools

Dr. C. Rhodes (Recorder and Consultant to the Committee), Alberta Education

Dr. H. G. Sherk was an ex officio member of all MACOSA project committees.